

A Grammar of Sakhalin Ainu

enciw itah

Aynu Itah

itah.aynu.org · 2026-06-22

Contents

Part I · The language and its setting	1
Introduction	1
Chapter 1 · The language and its speakers	3
Chapter 2 · Dialects, sources, and previous research	8
Chapter 3 · Orthography and transcription	15
Part II · Phonology	21
Chapter 4 · Phonology	21
Chapter 5 · Morphophonology	28
Part III · Word classes and the noun phrase	34
Chapter 6 · Word classes	34
Chapter 7 · Nouns and possession	39
Chapter 8 · Number	44
Chapter 9 · Pronouns and demonstratives	50
Chapter 10 · Numerals and quantification	54
Chapter 11 · Postpositions and relational nouns	60
Part IV · The verb	66
Chapter 12 · Verb structure and transitivity	66
Chapter 13 · Person marking and alignment	71
Chapter 14 · Valency-changing morphology	78
Chapter 15 · Tense, aspect, and modality	83
Chapter 16 · Evidentiality	89
Chapter 17 · Negation	95
Part V · Syntax	102
Chapter 18 · Copula, existence, and possession clauses	102
Chapter 19 · The simple clause and word order	104
Chapter 20 · Nominalization	107
Chapter 21 · Relative clauses	110
Chapter 22 · Complementation	113
Chapter 23 · Clause linking	116
Chapter 24 · Information structure	121
Part VI · The lexicon and language contact	126
Chapter 25 · Language contact and loanwords	126
Part VII · Texts	131
Chapter 26 · Glossed texts	131
Appendices	140
Abbreviations and conventions	140
References	141

Part I · The language and its setting

Introduction

What this grammar sets out to do, the data and conventions it relies on, how to read its glossed examples, and the dialect labels and scholarship behind them.

This book is a reference grammar of Sakhalin Ainu, the branch of the Ainu language family spoken on the southern half of Sakhalin until the late twentieth century and documented in a record that runs from the 1860s to the 1990s. It is organised to be consulted rather than read straight through: six parts lead from the speakers and their history through phonology and morphology to the syntax of the complex clause, and every numbered example is glossed, sourced, and tagged for dialect. This chapter explains how the book works — what it sets out to do (§0.1), what data it admits and under what conventions (§0.2), how a numbered example is put together (§0.3), what the dialect labels mean and which variety is the default (§0.4), and whose scholarship the description rests on (§0.5).

0.1 Aims and scope

Sakhalin Ainu has been described in parts, but never whole. Chiri planned a complete grammar and finished only the phonology and a portion of the morphology (Chiri 1942); Murasaki described the sound system and morphology of a single west-coast variety, Rayciska (Murasaki 1979); the recent literature treats individual subsystems — evidentiality, negation, person marking, number — one article or dissertation at a time. The aim of this book is the obvious missing item: a single, systematically organised description of the language from phoneme to discourse, written in English, in which each statement can be traced to the evidence for it.

Two consequences of that aim shape the text. First, the grammar is a synthesis where a synthesis is possible: points established by earlier scholars are credited to them by name and page, and the reader should be able to move from any paragraph here to the passage of Chiri, Murasaki, Piłsudski, or Dal Corso that stands behind it. Second, where the literature is silent — and for much of the syntax it is — the grammar does not merely record the silence. It analyses the texts directly and states its own conclusion, marked as such, so that a gap in previous research becomes a description that later work can confirm or correct. The data underneath such analyses remain attested; only the analysis is new.

The grammar treats Sakhalin Ainu as a system in its own right, not as a divergent dialect of the better-described Hokkaidō varieties. No knowledge of Hokkaidō Ainu is assumed, and Hokkaidō forms appear only where an explicit comparison sharpens the Sakhalin picture — most often in phonology (Chapter 4), where the correspondences are systematic, and in the verb, where the two branches have grammaticalised differently (Chapter 13, Chapter 16).

0.2 Data and conventions

Every numbered example in this grammar is attested unless it carries the explicit label *constructed example*. Constructed material is used sparingly — to complete a paradigm table or to spell out the prediction of an analysis first made here — and is never the sole support for a descriptive claim. The attested examples come from the published record of the language: Piłsudski's *Materials* (Piłsudski 1912), the texts embedded in Dobrotvorsky's dictionary (Dobrotvorsky 1875) in Sakaguchi's edition (Sakaguchi 2021), the letters of Sentoku Tarōji (Tangiku & Ogihara 2001), the folktales of Asai Take (Murasaki 2001), the post-war recordings and elicited sentences of Hattori and Murasaki, and the examples of the grammatical literature itself. The full inventory, with an assessment of what each source can and cannot show, is given in Chapter 2.

Examples are glossed morpheme by morpheme after the Leipzig Glossing Rules (Comrie, Haspelmath & Bickel 2015). Each grammatical gloss is set in small capitals and expands on hover; the complete list, with the handful

of additions this grammar makes for Sakhalin-specific categories, is kept on the abbreviations page, and the build of this site checks every gloss in every example against that list, so the two cannot drift apart.

Citation runs in an unbroken chain. A statement in the prose is supported by a numbered example; the example cites its source down to the page, text, or letter; author–year citations such as Chiri (1942: §76) link to the references, where works that could only be consulted second-hand are flagged as reported. In the other direction, every morpheme in an example, and every italicised Ainu form in the prose such as *cise* ‘house’, links to its entry in the dictionary of this site, which in turn indexes the corpus attestations of the word. The reader who distrusts a claim is meant to be able to reach the primary record without leaving the page’s links.

0.3 How to read an example

The following, from a Tarayka tale in Pilsudski’s collection, shows all the parts.

(1) *Emújke rajaxcí.*

emuyke *ray-ahci*
all die-PL

‘They all died.’ — Pilsudski 1912: text 2, l. 4; East Sakhalin, Tarayka (Sisratoka)

The number in parentheses counts examples within the chapter; cross-references in the prose use it. The first line, in italics, is the sentence as the source prints it — here Pilsudski’s own transcription, with his accent marks and his *x* for the velar realisation of the coda we write *h*. When source spelling and our romanisation coincide, this line is omitted. The second line is the morphemic segmentation in the romanisation of this grammar (Chapter 3): a hyphen separates affixes, as in *ray-ahci* ‘die’ plus the plural suffix, and an equals sign marks clitics. Each piece links to its dictionary entry.

Beneath it, the gloss line gives one gloss per word: lexical glosses in lower case (‘all’, ‘die’), grammatical categories in small capitals (PL), each of the latter expandable by hovering. Then comes the free translation in single quotes, sometimes followed by a literal rendering marked *lit.* where the idiom and the morphology pull apart. The last line is the source: the citation (Pilsudski 1912, text 2, line 4), the dialect tag (§0.4), and, where known, the locality and the speaker — here Sisratoka of Tarayka. Some examples close with a small-print note on points the gloss cannot carry, such as the note on *-ahci* spelt *-axcí* in Chapter 26, where this tale is glossed in full.

0.4 Dialect labels and the West-coast baseline

Sakhalin Ainu was not uniform: the record splits, broadly, into a west-coast and an east-coast dialect group, surveyed in Chapter 2. Every example in this grammar carries one of three tags.

Tag	Variety	Chief sources
WS	West Sakhalin: Rayciska, Odasu, Maoka, Taran-tomari, and neighbours	Murasaki’s and Hattori’s recordings; the Asai Take folktales
ES	East Sakhalin: Tarayka, Otasan, Ay, Tunayci, and neighbours	Pilsudski’s <i>Materials</i> ; the Sentoku letters
SA	Sakhalin, not further localisable	early wordlists; generalising statements in the literature

Table 1: Dialect tags used on examples

Unmarked descriptive statements in the prose describe the west-coast varieties. This is a policy, not a judgement of importance: the west coast is where the twentieth-century record is deepest, with sound recordings, elicitation, and a described phonology, so it is the only variety for which negative evidence — the claim that something does *not* occur — carries weight. East-coast data are flagged ES wherever they diverge from the baseline or supply what the west-coast record lacks; since no descriptive grammar of any east-coast dialect exists, Pilsudski’s texts and Sentoku’s letters must serve there in place of recordings. A WS example for comparison with the ES example above:

- (2) *mukara ani sitayki ike mukara-ha kotahma*
 axe INS hit and axe-POSS stick

‘He struck it with an axe, and his axe stuck fast.’ — Murasaki 2001: text 35; West Sakhalin, Odasu (Asai Take)

Hokkaidō Ainu is not a dialect of this grammar’s object language and receives no tag; where a Hokkaidō form is cited for contrast, the prose says so explicitly, and the references page marks Hokkaidō-centred works as such. Nothing from Hokkaidō is ever silently folded into the description of Sakhalin.

0.5 The scholarship behind this grammar

The documentation of Sakhalin Ainu is the work of a short and traceable line of scholars, and this grammar is best read as the latest link in it. The first substantial record is Dobrotvorsky’s dictionary of 1875, compiled on the west coast while Sakhalin was a Russian penal colony (Dobrotvorsky 1875). The largest body of connected text is owed to Bronisław Piłsudski, whose *Materials* of 1912 transcribe and translate east-coast narratives collected in 1902–1905 (Piłsudski 1912). Chiri Mashihō then gave the language its first linguistic grammar, the 1942 アイヌ語法研究, written with the Sakhalin dialect at its centre though never carried as far as syntax (Chiri 1942).

After the war, Hattori Shirō brought the methods of structural phonology and dialect geography to the surviving speakers, producing the comparative dialect dictionary and the first systematic account of Sakhalin person marking (Hattori 1964), (Hattori 1961). From the 1960s Murasaki Kyōko recorded the last fluent generation on the west coast — above all Asai Take, the final native speaker, who died in 1994 — and published the materials, the grammar, and the folktale corpus that anchor the modern study of the language (Murasaki 1976), (Murasaki 1979), (Murasaki 2001).

The present generation has re-edited the old records and pushed the analysis into territory the documenters never treated: Anna Bugaeva on nominalisation and Ainu typology at large (Bugaeva 2016), (Bugaeva 2022), Elia Dal Corso on evidentiality, alignment, and the Piłsudski and Murasaki corpora (Dal Corso 2018), (Dal Corso 2024), Sakaguchi Ryō on number, numerals, and the philology of the early sources (Sakaguchi 2021), (Sakaguchi 2024), and Tangiku Itsuji on the Sentoku letters and the Sakhalin–Hokkaidō divide (Tangiku & Ogihara 2001), (Tangiku 2022). The full history of this research, source by source, is the subject of Chapter 2.

Chapter 1 · The language and its speakers

Sakhalin Ainu and its names, its speakers and their history, language contact, endangerment, and documentation.

Sakhalin Ainu is the variety of Ainu once spoken across the southern half of Sakhalin island. Together with Hokkaidō Ainu and Kuril Ainu it forms one of the three major branches of the small Ainu language family (Janhunen 2022). The language ceased to be passed on to children in the first half of the twentieth century and lost its last native speaker in 1994 (Dal Corso 2024: 55, 58), so that everything we know of it rests on a documentary record stretching from the 1860s to the 1990s. This chapter introduces the language and its speakers; Chapter 2 surveys the dialects and the sources in detail.

1.1 Names of the language and its speakers



A Sakhalin Ainu man in a traditional robe with appliqué designs and a woven belt. Bronisław Piłsudski, c. 1905 · public domain · Wikimedia Commons

The name *Ainu* is the word *aynu* ‘human being, person’, which in opposition to other ethnonyms and to the gods denotes members of the Ainu people; the word was in use in this sense among all regional groups, including those of Sakhalin (Janhunen 2022), (Hattori 1964: 34). Sakhalin speakers applied it to themselves in ordinary discourse: writing to Bronisław Piłsudski in 1906, Sentoku Tarōji 千徳太郎治 — himself a Sakhalin Ainu, who would later publish 樺太アイヌ叢話, one of the earliest accounts of the community written from within it (Sentoku 1929) — reports on his community as follows.

(1) *anoakai aynu utara ramma ka pirka*

<i>anoakai</i>	<i>aynu</i>	<i>utara</i>	<i>ramma</i>	<i>ka</i>	<i>pirka</i>
we	Ainu	PL	always	even	be.well

‘we Ainu people are as well as ever’ — Tangiku & Ogihara 2001: Letter 1; East Sakhalin, Otasan (Sentoku Tarōji)

Alongside *aynu*, Sakhalin Ainu has a second word for ‘human being’ with no counterpart in Hokkaidō: *enciw* ‘human being’, recorded by Chiri (1942: §2) in the form *enciw* among words showing voicing after a nasal, and used historically as a self-designation by the Sakhalin Ainu (Dal Corso 2024: 55). The modern descendants of the Sakhalin Ainu have adopted *enciw* as the label that distinguishes them from the Hokkaidō Ainu (Janhunen 2022), and in today’s revival movement the language is called *Enciw itah* ‘language of the Enciw’ (Dal Corso 2024: 55), a usage also followed by (Murasaki 2025).

The second element of that name, *itah* ‘speech, language’, is the Sakhalin form corresponding to Hokkaidō *itak*, with the regular Sakhalin coda neutralisation described in Chapter 4. It is an old element of the language’s own metalinguistic vocabulary: the title page of the first dictionary of the language renders ‘Ainu–Russian dictionary’ in Cyrillic-script Ainu as айну-русский итаку-чóмень, with *itaku* for ‘word, speech’ (Dobrotvorsky 1875). In the scholarly literature the variety is most often called Sakhalin Ainu in English and 樺太アイヌ語 *Karafuto ainugo* ‘Karafuto Ainu’ in Japanese, after the Japanese name of the island (Dal Corso 2024: 57); both names are used of the same object, and this grammar writes Sakhalin Ainu.

1.2 Geographic setting and history

The Ainu settlement area on Sakhalin covered the southern part of the island, from Aniwa Bay in the south up both coasts to roughly the 50th parallel. Archaeological evidence places the Ainu expansion from Hokkaidō onto southern Sakhalin at about 1300 CE (Janhunen 2022). On the west coast, facing the Strait of Tartary, lay villages such as Tarantomari 多蘭泊 (now Калинино), Maoka 真岡 (now Kholmsk), Rayciska 来知志 (now Красногорск), Odasu 小田洲 (now Парусное), Usoro 鵜城, and Esutoru 恵須取 (now Углегорск); on the

east coast, facing the Sea of Okhotsk, Ociho 落帆, Tunayci 富内, Ay, Otasan 小田寒, Hunup, and, on the Bay of Patience in the north-east, Tarayka 多来加 and Nairo 内路 (Tangiku 2022), (Ono 2020: 232), (Dal Corso 2024: 58). The Poronay basin behind Tarayka was the zone where the Ainu were in community-level contact with the Uilta and the Nivkh, their only ethnic neighbours on the island apart from Japanese and Russians (Janhunnen 2022) (see §1.4).



The Sakhalin Ainu settlement world (schematic). The West coast faces the Strait of Tartary and the East coast the Sea of Okhotsk; Hokkaidō lies across the La Pérouse (Sōya) Strait to the south.



A Karafuto (Sakhalin) Ainu dwelling and its occupants; the signboard reads 樺太アイヌ家 'Karafuto Ainu House'. Torii Ryūzō, 1912 · public domain (PD-Japan) · Wikimedia Commons

Both colonial powers reshaped Ainu life on Sakhalin in the nineteenth century. Japanese fisheries operated along the coasts, with Ainu working in and later running them (Pilsudski 1912: X), while Russia ran the island as a penal colony; it was as a military physician in this colony, in 1867–1872, that Mikhail M. Dobrotvorskij Михаил Михайлович Добровровский made the first substantial record of the language (Sakaguchi 2021: 43). By the Treaty of Saint Petersburg of 1875 the whole island passed to Russia, and 841 Ainu of the Aniwa Bay area were moved to Sōya in northernmost Hokkaidō and then, against their declared wishes, resettled by the colonial administration at Tsuishikari 対雁 near Sapporo. Cholera and smallpox epidemics in 1879 and 1886–1887 killed about half of them; survivors began returning to Sakhalin around 1890, and after the island's southern half became the Japanese territory of Karafuto 樺太 in 1905 nearly all of the Tsuishikari community went home (Sakaguchi 2024: 7–8).

At the start of the Karafuto period (1905–1945) the Sakhalin Ainu numbered about 2,000, and nearly all of them spoke the language (Tangiku 2022); Pilsudski (1912: V) put the entire Ainu population of Hokkaidō, Sakhalin, and Shikotan at about 20,000 at the same date. Japanese assimilation policy discouraged the use of Ainu and pressed a Japanese lifestyle on the community, and transmission to children broke off in the first half of the twentieth century (Dal Corso 2024: 57). Small, mainly Ainu hamlets nevertheless persisted on Karafuto until the end of the Second World War, and several hundred speakers are estimated to have remained in 1945 (Tangiku 2022). When southern Sakhalin reverted to the Soviet Union after the war, almost the whole Ainu population, holding Japanese citizenship, was evacuated to Hokkaidō and points south by 1948 and dispersed there; only a small number stayed behind in Russia (Sakaguchi 2024: 8), (Dal Corso 2024: 57).

The fluent speakers known to post-war research were all evacuees living in Hokkaidō. The two whose speech is best documented are Fujiyama Haru 藤山ハル (1900–1974), raised in Esutoru and long resident in Rayciska, who was found through the dialect survey of the 1950s living at Tokoro in eastern Hokkaidō (Tangiku 2022), (Sakaguchi 2024: 8), and Asai Take 浅井タケ (1902–1994) of Odasu (Tangiku 2022). With Asai Take's death in 1994 the language lost its last native speaker (Dal Corso 2024: 55, 58).

1.3 Genetic affiliation and typological profile

Ainu has no demonstrated relatives. Proposals linking it to “Altaic”, Japanese, Nivkh, and other languages have not withstood scrutiny (Shibatani 1990: 5–7), (Vovin 1993), so Ainu is either an isolate or, given its considerable internal diversity, a small family of three closely related languages corresponding to the Hokkaidō, Sakhalin, and Northern Kuril groups (Janhunen 2022). The position of Sakhalin Ainu as a primary branch is supported quantitatively: in the lexicostatistical survey conducted by Hattori Shirō 服部四郎 and Chiri Mashiho 知里真志保 in the 1950s, the cluster analysis of Asai (1974: 100) set the Sakhalin dialects apart from both Hokkaidō and North Kuril as one of three top-level groups, and the statistical reanalysis of Ono (2020) yields an even

sharper picture in which the deepest split in the family runs between the Sakhalin dialects and everything else. A Bayesian phylogeographic analysis of the same Hattori–Chiri wordlists – which sample the Hokkaidō and Sakhalin dialects but not Kuril – likewise splits them into a Sakhalin and a Hokkaidō subgroup and traces their spread from a homeland in northern Hokkaidō (Lee & Hasegawa 2013). Sakhalin Ainu separates from Hokkaidō Ainu by regular phonological correspondences and by its own grammatical retentions and innovations; Vovin (1993) draws on Sakhalin evidence, including the Nairo dialect, throughout his reconstruction of Proto-Ainu. The internal division of Sakhalin Ainu is treated in Chapter 2.

Typologically, Sakhalin Ainu is a polysynthetic, agglutinating, strongly head-marking language with verb-final order: intransitive clauses are SV and transitive clauses AOV (Dal Corso 2024: 56). Adpositions follow their complement (Chapter 11), and modifiers precede their head. Personal prefixes and clitics on the verb index the core arguments, in a mixed alignment system that includes the indefinite set *an-*, *-an*, *i-* discussed in Chapter 13 (Dal Corso 2025a). Nouns distinguish a plain and a possessed (“belonging”) form, and the possessive suffix has a wider range in Sakhalin than in Hokkaidō, extending into the nominalisation of verbs (Chapter 7, Chapter 20) (Bugaeva 2016), (Tangiku 2022). The phonology has five vowels with a length contrast and a coda system in which the stops of Hokkaidō Ainu are neutralised to *h* (Chapter 4) (Tangiku 2022), (Itabashi 2001). There is no tense morphology; mood, aspect, and evidentiality are expressed by synthetic and analytic means (Dal Corso 2024: 56), among them the reportative *manu* that recurs at the end of narrative sentences throughout the folktale corpus (Chapter 16) (Dal Corso 2018):

- (2) *orowa sine monimahpo neampe poniwne monimahpo nee manu*
 then one girl TOP be.younger girl COP REP

‘and one of the girls was the younger daughter, it is said’ — Murasaki 2001: text 19; West Sakhalin, Odasu (Asai Take)

Numerals are built on a base of twenty, *hoh* ‘twenty’, with vigesimal counting amply attested in the nineteenth-century record; alongside it the Sakhalin sources show borrowed decimal bases such as *tanku* ‘hundred’ and *wantanku* ‘thousand’, and decimal counting gained ground over the period of documentation (Chapter 10) (Sakaguchi 2022), (Murasaki 2009). Negation is analytic, with preverbal negators such as *hannehka* that developed within Sakhalin Ainu (Chapter 17) (Dal Corso 2021), (Dal Corso 2025b).

1.4 Language contact

On Sakhalin the Ainu bordered on the Uilta (Orok), speakers of a Tungusic language, and the Nivkh, speakers of another isolate; the principal contact area was the Poronay basin and the Tarayka region in the north-east, with a second mixed zone on the northern west coast (Janhunnen 2022), (Shiraishi & Tangiku 2022), (Gruzdeva 1996). Japanese travel records of the late eighteenth and nineteenth centuries show that multilingualism was limited and localised rather than general: in contact situations around Lake Tarayka it was usually the Uilta who spoke Ainu, while command of Ainu among the Nivkh was restricted to particular individuals (Shiraishi & Tangiku 2022).

Lexical borrowing nevertheless ran in both directions. Sakhalin Ainu took over dog-sled vocabulary from Nivkh, including *nuso* ‘dog sled’ (also *nuhsu*) from Nivkh *nuci* – an etymology going back to Chiri – and shows pairs such as *muhru* ‘pillow’ beside Nivkh *mut* with the regular *r*-for-*t* correspondence of loans between the two languages (Shiraishi & Tangiku 2022). From Uilta the west-coast dialects took *wampakka* ‘mittens’, where the east coast uses *matumere* (Shiraishi & Tangiku 2022). Older strata of contact include *tunakay* ‘reindeer’, connected with the Nivkh (Ghilyak) word for the pulling animal, and the early Japonic loans *sippo* ‘salt’ and *pasuy* ‘chopsticks’, which entered Ainu as a whole well before the historical period (Janhunnen 2022). The decimal numeral bases mentioned in §1.3 are likewise borrowings (Sakaguchi 2022: 67–68).

Japanese supplied the names of trade goods that reached the Ainu from the south, for example *hasami* ‘scissors’ and *konkaani* ‘gold, metal’ (Shiraishi & Tangiku 2022), and during the Karafuto period it became the dominant contact language altogether; the post-war recordings of Asai Take contain frequent Japanese asides and code-switches around an Ainu grammatical core (Murasaki 2001). Russian left a lighter mark on the lexicon, but a distinctive one on writing: pupils of the school Pilsudski ran on the east coast learned to write their own language in Russian characters (Pilsudski 1912: XII), and Sentoku Tarōji’s Cyrillic-script letters of 1906 – see

(1) – are the chief monument of this practice (Tangiku & Ogihara 2001) (on script, see Chapter 3). The Russians themselves appear in the letters under the Ainu name *nuca* ‘Russian’.

1.5 Endangerment and present-day status

Sakhalin Ainu is today a dormant language: natural transmission ended in the first half of the twentieth century, and no native speaker has lived since 1994 (Dal Corso 2024: 55, 57), (Janhunen 2022); the UNESCO *Atlas of the World’s Languages in Danger* accordingly records Sakhalin Ainu among the languages that have become extinct since they were linguistically described (UNESCO 2010). It is not, however, an unused one. Descendants of the Sakhalin Ainu in Japan study and revive the language under the name *Enciw itah*, with the scholar and activist Kitahara Jirōta Mokottunas 北原次郎太 among those leading the effort (Dal Corso 2024: 55), and teaching materials now exist, including the conversation course of Murasaki (2025). The recorded legacy is substantial: Piłsudski’s wax-cylinder recordings of the Ainu, made in 1902–1903, have been partly restored from the 1980s onward by a laser-beam reflection method developed for the purpose (Iwai et al. 1987), (Dal Corso 2024: 59), and the post-war tape recordings of Fujiyama Haru and Asai Take were published with audio (Murasaki 2001), (Tangiku 2022).

This grammar is one part of that documentation-based revival. It is published beside a dictionary of Sakhalin Ainu on this site, and every Ainu word and morpheme in the examples is linked to the dictionary so that the attested record, rather than any modern reconstruction, remains the point of reference throughout. The sources on which the description rests are set out in Chapter 2.

Chapter 2 · Dialects, sources, and previous research

The West and East coast dialect groups, the major text collections and recordings, and earlier grammatical descriptions.

Sakhalin Ainu is known from a record built up in four waves: the Russian-era documentation of the 1860s–1900s (Dobrotvorsky, Piłsudski), the writings of native authors and Japanese scholars in the Karafuto period (Sentoku, Kindaichi, Chiri), the post-war work with evacuee speakers in Hokkaidō (Hattori, Murasaki, Tamura), and the philological and descriptive research of recent decades. The waves do not sample the language evenly: the nineteenth- and early twentieth-century sources come mostly from the east coast, the post-war recordings almost entirely from the west coast. This chapter first sets out the dialect landscape and then surveys the sources and the research built on them.

West coast (wsa) East coast (esa) · tap a marker for its names; ⌘/Ctrl-scroll to zoom

Place names West coast (wsa)

- Honto· 本斗 (ほんとう)· Nevelsk— South-west port on the approach to Aniva Bay.
- Tarantomari· 多蘭泊 (タラントマリ)· Kalinino— SW-coast settlement near Honto; West-coast narrative source.
- Mawka· 真岡 (マウカ)· Kholmsk— Principal west-coast port; one of Murasaki’s two reference dialects.
- Tomari-oro· 泊居 (トマリオロ)· Tomari— West-coast port; Ainu tomari-oro ‘within the harbour’.
- Kus-un-nay· 久春内 (クシュンナイ)· Ilyinsky— West-central area documented by Dobrotvorskij (1875).
- Rayciska· 来知志 (ライチシカ)· near Krasnogorsk (Chinnai)— Home dialect of Murasaki’s consultants Fujiyama Haru & Asai Take.
- Usoro· 鵜城 (ウソロ)· Orlovo— Northern west-coast village; Ainu usoro ‘bay, inlet’.
- Esturi· 恵須取 (エシトウリ)· Uglegorsk— North-central west-coast town; Fujiyama Haru’s birthplace.

East coast (esa)

- Kusunkotan· 大泊 (クシュンコタン)· Korsakov— Aniva Bay port; old Ainu Kusunkotan, poro-an-tomari ‘big harbour’.
- Tunayci· 富内 (トンナイチャ)· Okhotskoye— Southern end of Pilsudski’s narrator chain, at the outflow of Lake Tunaycha.
- Ocohoka· 落帆 (オチヨポカ)· Okhotskoye area— Pilsudski narrator hamlet by Lake Tunay (Ociho 落帆).
- Ay· 相浜 (アイ)· near Starodubskoye— Pilsudski’s 1902–04 base; home of chief Bafunke and narrator Kimura Usarusmah.
- Takoye· Takoye (near Dolinsk)— Pilsudski tuita narrator site in the Naiba valley.
- Otasam· 小田寒 (オタサム)· near Firsovo— SE-coast narrator hamlet on the Otasamu R.; Sentoku Tarōji’s home.
- Siraroko· 白浦 (シラウラ)· Vzmorye— Pilsudski narrator hamlet on the coast north of Dolinsk.
- Naayro· 内路 (ナイロ)· Gastello— East-coast village south of Shisuka; drawn on in Vovin’s reconstruction.
- Siisuka· 敷香 (シスカ)· Poronaysk— East-coast town at the Poronay mouth; hub of the Taraika region.
- Otasu· Poronaysk (delta)— 1920s multi-ethnic settlement (Uilta, Nivkh, Ainu) on the Poronay delta.
- Tarayka· 多来加 (タライカ)· Lake Nevskoye— The most conservative East-coast dialect; lagoon on the Gulf of Patience.

Figure 1: The Sakhalin Ainu dialect world. West-coast (wsa) settlements are shown in green, East-coast (esa) in amber; tap a marker for a place’s Ainu, Japanese (kanji), and present-day Russian names. The post-war recordings cluster on the west coast, Pilsudski’s folktale texts on the east.

2.1 Dialect division

The classical picture of Ainu dialect relations rests on the 200-item lexical survey carried out from 1955 by Hattori Shirō 服部四郎 and Chiri Mashiho 知里真志保 among evacuee speakers in Hokkaidō, which covered six Sakhalin localities: Ociho 落帆, Tarantomari 多蘭泊, Maoka 真岡, Shiraura 白浦, Rayciska 来知志, and Nairo 内路 (Sakaguchi 2024: 8), (Ono 2020: 231–232). The cluster analysis of Asai (1974: 100) grouped these six against all Hokkaidō dialects and North Kuril as one of three top-level branches, and divided them internally into a northern group (Nairo, in the Tarayka region), a central group (Ociho, Maoka, Shiraura, Rayciska), and a southern group (Tarantomari). Ono (2020) reanalysed Asai’s cognacy matrix with modern clustering methods and found the deepest division in Ainu to run between the Sakhalin dialects and everything else; the coherence of Sakhalin as a unit is robust under either analysis.



Two Sakhalin Ainu men — the one at left from the western coast, the one at right from the eastern. The original print is inscribed “Ainu of Saghalien.” Bronisław Piłsudski, c. 1902–05 · Smithsonian NAA · public domain · Wikimedia Commons

Within Sakhalin, the documentation sorts itself by coast. West-coast varieties are known from post-war speakers: Rayciska and Esutoru 恵須取 (Fujiyama Haru 藤山ハル), Odasu 小田洲 (Asai Take 浅井タケ), Usoro 鵜城 (the speaker recorded by Wada Bunjirō 和田文治郎), Maoka and Tarantomari (speakers recorded by Murasaki and by the 1950s surveys) (Tangiku 2022), (Dal Corso 2024: 57–58). East-coast varieties are known chiefly from the Russian era and the first Karafuto years: Ay, Tunayci 富内, Hunup, and Tarayka through Piłsudski’s texts, Otasan 小田寒 through Sentoku Tarōji’s letters, and Tunayci and Ociho through Kindaichi’s 1907 phonetic survey (Dal Corso 2024: 58), (Dal Corso 2025a), (Sakaguchi 2024: 8).

Tarayka and neighbouring Nairo in the north-east stand apart from the rest of Sakhalin. They retain the syllable-final stops *p t k* that other Sakhalin dialects neutralise to *h* (Tarayka *tek* ‘hand’, *sik* ‘eye’ against west-coast *teh*, *sis*), and they turn word-initial *r* into a stop in a number of words (Nairo *tup* ‘ice’ where Esutoru has *ruh*) (Tangiku 2022); Vovin (1993) treats this Nairo *t-* as continuing a distinct Proto-Ainu initial, his **d-*. In Asai’s matrix Nairo shares only 60–74 of 110 test items with the other Sakhalin localities, which themselves share 84–93 with one another (Asai 1974: 92), (Ono 2020: 234). The divergence matches the ethnographic position of the Tarayka people, who lived in the contact zone with the Uilta and Nivkh and are singled out as culturally distinct in the nineteenth-century Japanese records (Shiraishi & Tangiku 2022).

2.1.1 What differs between the coasts The two coastal groups differ at every level of structure; here we point to the differences taken up in later chapters. In phonology, beside the coda facts just mentioned (Chapter 4), the Piłsudski corpus shows open-mid values of *e* and *o* for the east coast where the west-coast recordings show close-mid vowels, frequent intervocalic voicing of plosives that is marginal in the west-coast materials, and no trace of the glottal stop heard in west-coast speech (Dal Corso 2024: 59–61). In the lexicon there are coast-bound pairs such as west-coast *wampakka* ‘mittens’ (a Uilta loan) against east-coast *matumere* (Shiraishi & Tangiku 2022), and finer-grained isoglosses within each coast, for example *rayki* ‘kill’ in Shiraura and Esutoru against *rayke* in Maoka, Tarantomari, and Ociho (Tangiku 2022).

The sharpest morphological difference found so far concerns person marking (Chapter 13). The east-coast dialects of Piłsudski’s corpus have a dedicated exclusive first-person plural set — *ci-* for transitive subjects, *-as* for intransitive subjects, and an object prefix *in-* found nowhere else in Ainu — while the west-coast materials lack the productive set, retaining at most marginal traces (Chapter 13 (Person marking and alignment)) (Dal Corso 2024: 66–68):

(1) *Kamúi utara inránu.*

kamuy utara in-ranu
 god PL 1PL.EXCL.O-love

‘the gods love me’ — Pilsudski 1912: 86; Dal Corso 2024: 67; East Sakhalin

in- is the exclusive first-person plural object prefix, here referring to the narrator, as is usual in folklore; the stem is *ranup* with elision of the final stop.

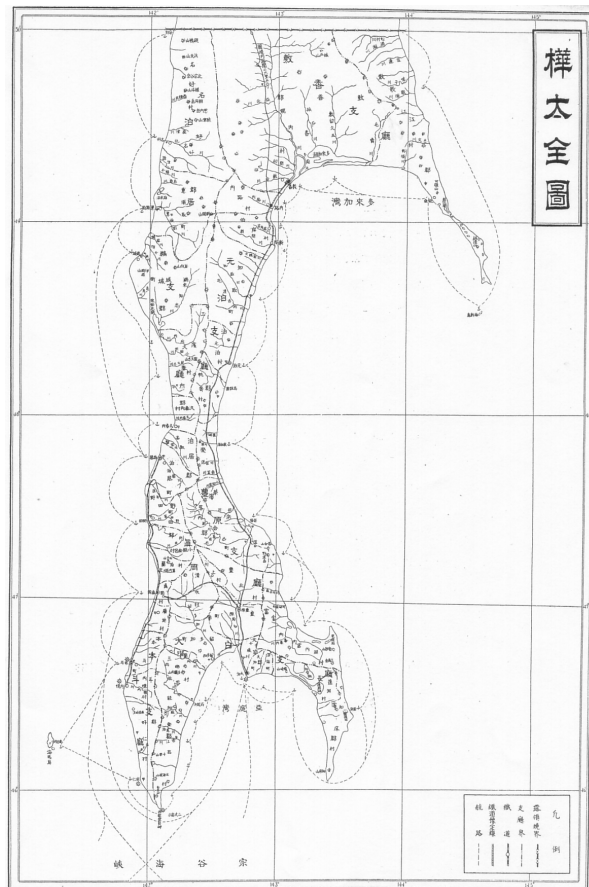
In the west-coast corpus the same job falls to the indefinite set: speakers refer to themselves with *an=* and its congeners, as in Asai Take’s

(2) *yee wa an=nuu*
 say and INDEF.A=hear

‘tell it, and I will listen’ — Murasaki 2001: text 5; West Sakhalin, Odasu (Asai Take)

Verbal morphology also diverges in the functions of the applicative and reflexive prefixes *e-*, *ko-*, and *yay-*, which show valency-preserving uses in the east-coast texts beyond what the west-coast materials attest (Chapter 14) (Dal Corso 2024: 68–76). Since the east-coast record predates the west-coast record by sixty to seventy years, each such difference must be weighed as possibly diachronic rather than purely geographic (Dal Corso 2024: 53–54).

2.2 The documentary record



Karafuto zenzu (樺太全圖), a 1930s Japanese map of Karafuto — the Japanese-administered southern half of Sakhalin under whose rule the Ainu documented here lived. Tokai-Tsushinsha, 1930s · public domain (PD-Japan) · Wikimedia Commons

Dobrotvorsky 1875. The Ainu–Russian dictionary of Mikhail M. Dobrotvorsky Михаил Михайлович Добротворский (1836–1874), compiled during his five years as a military physician on Sakhalin (1867–1872) and published posthumously in Kazan, is the earliest substantial record of the language (Dobrotvorsky 1875), (Sakaguchi 2021: 43). It numbers 10,930 entries, of which 5,733 — among them about 100 personal names

and 511 place names — come from his own fieldwork, the rest from earlier glossaries; his own material is recognisable by its accent marks (Sakaguchi 2021: 43). An appendix contains two connected texts: a speech addressed by Ciwokanke to the priest Simeon at Kusunay, and a reconciliation dialogue. The named consultants of the example sentences — Sambaku, Kasituru, Sukenno — lived on the central west coast, and the language of the examples, with first-person singular *cokay*, differs from the north-west-coast speech recorded after the war (Sakaguchi 2021: 43–44). The transcription is good for its date but runs together *i/e* and *o/u* and does not mark vowel length; the texts were re-edited, romanised, and translated by Sakaguchi (2021), whose edition this grammar follows when citing them.



Bronisław Piłsudski (1866–1918), whose 1902–05 recordings are the foundation of the East Sakhalin record. anonymous, before 1918 · public domain · Wikimedia Commons

Piłsudski 1912. Bronisław Piłsudski (1866–1918), exiled to Sakhalin in 1887, worked among the east-coast Ainu in 1896 and again from 1902 to 1905, latterly for the Russian Academy of Sciences (Dal Corso 2024: 56–57). His *Materials for the Study of the Ainu Language and Folklore* publishes 27 *ucaskuma* ‘tradition, historical tale’ out of a collection he himself put at 350 texts (Piłsudski 1912: XXI). The texts were dictated by named speakers of four east-coast villages — Tarayka, Hunup, Ay, and Tunayci — each introduced in the notes to his or her first text: Sisratoka, aged 28, of Tarayka, Ipoxni, aged 32, of Hunup, and several tellers at Tunayci, among them the 88-year-old Yasinoste (Piłsudski 1912), (Dal Corso 2024: 58). Most were monolingual first-language speakers, so the corpus shows the language with almost no Japanese or Russian admixture — a state never again captured (Dal Corso 2024: 54, 58). Each text carries an English translation and copious grammatical, lexical, and ethnographic notes, and the preface includes Piłsudski’s description of his transcription, a close phonetic one which he was able to check with the phonetician Rousselot in London in 1910 (Piłsudski 1912: XV), (Dal Corso 2024: 59). Its limitations are those of its method: dictation is slow, interrupts narration, and invites unnatural repetition, as Piłsudski himself observed (Piłsudski 1912: IX), (Dal Corso 2024: 59). He also recorded part of his collection on phonograph wax cylinders, many now degraded beyond playback, though some audio has been recovered by laser reading (Dal Corso 2024: 59). Counting later publications, some forty of his Sakhalin texts are now in print with the Ainu original (Sakaguchi 2024: 8).

Sentoku Tarōji. The one substantial body of Sakhalin Ainu written by a native speaker comes from Sentoku Tarōji 千徳太郎治 (Ainu name Taronci), a Tsuishikari Ainu who returned to Otasan on the east coast and worked with Piłsudski (Sakaguchi 2024: 8). His letters to Piłsudski, written in 1906 in Cyrillic-script Ainu, were deciphered and published with transliteration and Japanese translation by Tangiku & Ogiwara (2001); they are

first-hand epistolary prose, a register otherwise absent from the corpus (Chapter 3 discusses their spelling). His Japanese-language book 樺太アイヌ叢話 *Karafuto ainu sōwa* (Sentoku 1929) contains two further oral-literature texts in Ainu (Sakaguchi 2024: 8).

The Karafuto period. Kindaichi Kyōsuke 金田一京助 visited the new territory in 1907 and published detailed reports on the phonetics of Tunayci and Ociho, followed by the autobiography of the Tsuishikari-raised speaker Yamabe Yasunosuke 山辺安之助, あいぬ物語 *Ainu monogatari* (1913), with a grammatical outline as its appendix — the first systematic sketch of the language (Sakaguchi 2024: 8). The central work of the period is Chiri (1942), on which see §2.3. Texts taken down by the physician Wada Bunjirō from a speaker of the Usoro dialect have been published in instalments by Kitahara (2013–2019) (Dal Corso 2024: 57–58), and folktales by Numahata Ume 沼端ウメ of Usoro have likewise appeared from manuscript (Sakaguchi 2024: 9).

Post-war fieldwork. After the evacuation, all fieldwork took place in Hokkaidō. The Hattori–Chiri survey of 1955 (published 1960) supplied the comparative lexicon analysed by Asai (1974) and led to Hattori’s dialect dictionary (Hattori 1964), whose Sakhalin columns rest largely on Rayciska forms from Fujiyama Haru; Hattori also published studies of her dialect’s age-graded vocabulary (1957), accent (1961), and vowel length (1967) (Sakaguchi 2024: 8–9), (Tangiku 2022). Murasaki Kyōko 村崎恭子 recorded Fujiyama Haru extensively: her text collection (Murasaki 1976) presents folktales and conversations of Fujiyama Haru and of Ōta Yuk, a speaker of the Maoka dialect, and her grammar (Murasaki 1979) is the only full descriptive grammar of any Sakhalin dialect; the 1976 corpus was re-edited with glosses and an English translation by Dal Corso (Dal Corso 2021a) (on which see also Dal Corso (2024: 57–58)). Murasaki’s earliest analysis of Rayciska syntax appeared under her maiden name as (Furukawa 1967).

The last speakers. Asai Take of Odasu was recorded by Murasaki from the mid-1980s into the early 1990s; the corpus files of the present project carry session dates of 1984–1991. A first volume of the folktales was issued privately in 1996 (Murasaki 1996); the complete set then appeared with audio as 改訂浅井タケ昔話全集 in two volumes in the ELPR series (Murasaki 2001), (Tangiku 2022). A further ELPR volume edited by Tamura Suzuko 田村すす子 presents the oral literature of Yamada Hayo 山田ハヨ, recorded in 1958, together with other Sakhalin sound materials (Kitahara et al. 2003). Murasaki has since published shorter collections of Fujiyama Haru material (2013, 2016) (Tangiku 2022) and a conversation course based on her field data (Murasaki 2025).

2.2.1 Generational variation The Rayciska record is deep enough to show variation in apparent time as well as in space. Hattori’s dictionary sets certain forms apart as *onne itah* ‘elders’ speech’, distinct from what his consultants themselves said: the score numeral *hohne* ‘twenty’ carries this label, beside the everyday decimal *tu kunkutu* (Hattori 1964: 262), (Murasaki 2009: 74), and the same consultants knew the exclusive pronoun *cioka* only as the speech of elders or of the southern west coast (Hattori 1961: 5), (Sakaguchi 2024: 27). Murasaki draws the same line in her course materials, listing elders’ *kaa* and *enkaa* beside everyday *kaske* and *enkaske* among the relational nouns (Murasaki 2025: 17–18). These are not mere memories of a lost stage: the doublets coexist in running speech — Asai Take’s fifth folktale has both *kasketa* and elders’ *kaata* ‘on top of’ within a single passage (Murasaki 2001: text 5) — though the elders’ member could be moribund, as when Naert’s Rayciska informant produced the score construction *wanpe ikasma sine hoh* ‘thirty’ only as an archaism (Sakaguchi 2022: 71). Hattori’s 1957 article on the age-graded vocabulary of Fujiyama Haru’s dialect, noted above, was the first study to take the phenomenon as its object (Sakaguchi 2024: 8–9).

	everyday	elders' (<i>onne itah</i>)	sources	discussion
‘twenty’	<i>tu kunkutu</i>	<i>hohne</i>	(Hattori 1964: 262); (Murasaki 2009: 74); (Murasaki 2025: 25)	Chapter 10
‘on top of (in contact)’	<i>kaske</i>	<i>kaa</i>	(Murasaki 2025: 17–18); (Murasaki 2001: text 5)	Chapter 11
‘above (detached)’	<i>enkaske</i>	<i>enkaa</i>	(Murasaki 2025: 17–18)	Chapter 11
‘we (exclusive)’	indefinite <i>an=</i> , <i>anoka</i>	<i>cioka</i> (or southern WS)	(Hattori 1961: 5); (Sakaguchi 2024: 27)	Chapter 13

Table 2: Elders’-speech doublets in the post-war Rayciska record. Each pair is analysed in the chapter cross-referenced at right.

This generational layering must be kept apart from the register and genre variation that runs through the same corpus. The alternation between the KU-series and the indefinite AN-series in first-person reference tracks text type, not speaker age: the Sentoku letters use *an=* twenty-three times and *ku=* never, Yamabe’s autobiography is narrated almost wholly in the AN-series, while Asai Take’s *tuytah* ‘folktale’ use *ku=* freely in dialogue (Sakaguchi 2024: 46) (Chapter 13). Both kinds of variation bear on the warning given in §2.1.1: with an east-coast record two generations older than the west-coast one, differences between the coasts may be differences between periods, and the elders’ doublets show such change caught in progress within a single community.

2.3 Previous research

The first extensive grammar centred on Sakhalin Ainu is アイヌ語法研究 *Ainu gohō kenkyū* by Chiri Mashiho (Chiri 1942), published in the reports of the Karafuto Government Museum in Toyohara. Chiri planned three parts — phonology, morphology, and phraseology — but completed only the phonology and part of the morphology, never reaching syntax, as his preface records (Chiri 1942: まへがき). The work treats Ainu as a whole with Sakhalin at the centre, systematically marking Hokkaidō and Karafuto forms, and its phonology establishes points still cited today: the five-vowel system, the Sakhalin-only coda *x* (our *h*), the exclusion of stops and *r* from Sakhalin codas, and the lengthening of accented vowels (Chiri 1942: §§1–6). Its examples are drawn partly from Piłsudski’s texts, partly from Chiri’s own Karafuto fieldwork (Sakaguchi 2024: 8).

Murasaki’s grammar (Murasaki 1979) describes the phonology and morphology of Rayciska from her own recordings and remains the reference description; Piłsudski’s grammatical notes, Kindaichi’s 1913 outline, and Hattori’s articles are its main predecessors (Sakaguchi 2024: 8–9). Specialised studies of the post-war corpus include Tangiku (1998) on west-coast verbs in *-rV*, Itabashi (2001) on the correspondence between Sakhalin vowel length and Hokkaidō pitch accent, and Murasaki (2009) on numerals.

Recent work falls into three strands. One is typological reassessment: Bugaeva (2016) argues that the Sakhalin possessive-based nominalisations are innovations, and the Handbook of the Ainu Language (Bugaeva 2022) contains chapters directly on Sakhalin — Tangiku (2022) on the differences between Karafuto and Hokkaidō dialects and Shiraishi & Tangiku (2022) on contact among the languages of Sakhalin. The second is the description of grammatical subsystems across the whole record: Dal Corso on evidentiality (Dal Corso 2018), the rise of analytic negation (Dal Corso 2021), (Dal Corso 2025b), and morphological alignment (Dal Corso 2025a); Sakaguchi Ryō 阪口諒 on *utara* (Sakaguchi 2020a), the nominal plural suffix *-ahcin* (Sakaguchi 2020b), numerals and calculation (Sakaguchi 2022), and person and number marking across all sources in his dissertation (Sakaguchi 2024). The third is philology: Sakaguchi’s editions of early texts, including the Dobrotvorsky texts (Sakaguchi 2021), Tangiku Itsuji’s 丹菊逸治 edition of the Sentoku letters (Tangiku & Ogihara 2001), and Dal Corso’s re-edition of the Murasaki corpus and analysis of Piłsudski’s corpus (Dal Corso 2024). No descriptive grammar of any east-coast dialect yet exists (Sakaguchi 2024: 9); for the east coast, Piłsudski’s corpus and the Sentoku letters must do the work that recordings do for the west.

2.4 The corpus used in this grammar

The examples in this grammar are drawn from the published record surveyed above: the Asai Take folktales (Murasaki 2001), the Sentoku letters (Tangiku & Ogihara 2001), the Dobrotvorsky texts in Sakaguchi's edition (Dobrotvorsky 1875), (Sakaguchi 2021), the Yamada Hayo recordings (Kitahara et al. 2003), Piłsudski's *Materials* (Piłsudski 1912), the examples in Chiri's grammar (Chiri 1942), and Murasaki's course sentences (Murasaki 2025).

Every example cites the original publication, with text, letter, page, or section number, and carries a dialect tag — WS for the west coast, ES for the east coast, SA where no finer localisation is possible — together with the locality and speaker where these are known. All Ainu material is retranscribed into the romanisation defined in Chapter 3; where a source's own spelling is itself the point, it is reproduced as an additional line, as with the Cyrillic of the Sentoku letters in Chapter 1.

Chapter 3 · Orthography and transcription

How Sakhalin Ainu has been written — in Latin, Cyrillic, and katakana — and the romanisation used in this grammar.

Sakhalin Ainu has no traditional script, but it has a longer history of being written than is often assumed, and in three different scripts. Russian observers and at least one native writer used Cyrillic between the 1860s and 1906; Bronisław Piłsudski and every later linguist used some form of the Latin alphabet; and from the Japanese period onward Ainu words and whole texts were set down in katakana, by researchers and by Ainu writers themselves. This chapter surveys these practices with attested specimens of each, then defines the romanisation used throughout this grammar and the Latin–Cyrillic–katakana correspondences used by this site. The phoneme inventory that the romanisation encodes is justified in the phonology chapter; the sources themselves are surveyed in the chapter on dialects and sources.

3.1 Cyrillic

Cyrillic is the script of two of the earliest bodies of connected Sakhalin Ainu text: the specimens appended to Mikhail M. Dobrotvorsky's (Михаил Михайлович Добротворский) dictionary of 1875, and the letters that Sentoku Tarōji (千徳太郎治) wrote to Piłsudski in 1906.

3.1.1 Dobrotvorsky and the earliest records Dobrotvorsky, a Russian military physician, lived on Sakhalin from 1867 to 1872 and compiled an Ainu–Russian dictionary of over ten thousand headwords, more than half of them collected by himself, mainly from speakers of the central West coast such as Sambaku of Nayoro Sakaguchi (2021: 43–44). At the end of the dictionary he printed two short connected pieces — Ciwokanke's speech and a reconciliation dialogue — which Piłsudski already singled out as the first usable Sakhalin texts (Piłsudski 1912: XXIV).

Dobrotvorsky wrote Ainu with the pre-reform Russian alphabet, marking stress but not vowel length (Sakaguchi 2021: 45–47). Several of his conventions are phonetically revealing. He alternates б/п and к/р for /p/ and /k/, recording the intervocalic voicing that is allophonic in Sakhalin Ainu (see phonology). Word-initial /r/ he consistently heard with an occlusive onset and wrote рп, as in рпéкопо for *reekoro* 'very' — an observation discussed at length by Piłsudski (1912: 8) and matched by Chiri's report that initial r wavers with a t-like stop (*reekoro* ~ *teekoro*) (Chiri 1942: 第1章). A silent ъ closes every consonant-final word, as in contemporary Russian spelling, and /c/ is normally ч. Example (1) shows the opening of his first text with our retranscription.

- (1) Танъ котанъ охтá утáса—анъ кусú áреги анъ.

tan kotan ohta utasa-an kusu ariki-an
this village in visit-INDEF.S PURP come.PL-INDEF.S

‘I have come to visit this village.’ — Dobrovorsky 1875: 90; Sakaguchi 2021: 49; West Sakhalin, West coast, central (recorded 1867–1872)

Dobrovorsky’s áреги shows both the voiced allophone of /k/ and the o~u, e~i wavering typical of unstressed vowels; the form is ariki ‘come.PL’, the plural stem regularly selected by the indefinite person.

3.1.2 Sentoku Tarōji’s letters of 1906 Sentoku Tarōji (1872–1929), an East-coast Ainu of Otasan schooled under the Russian administration, wrote three letters in Ainu to Pilsudski, dated 4 June, 15 June and 11 August 1906. They are kept in the Sakhalin Regional Museum and were deciphered and published with a Japanese translation by Tangiku Itsuji (丹菊逸治) and Ogihara Shinko (荻原真子) (Tangiku & Ogihara 2001). They are the earliest known Ainu-language texts written by an Ainu hand, antedating Chiri Yukie’s (知里幸恵) アイヌ神謡集 of 1923 by seventeen years (Nakagawa 2006: 21).

Sentoku’s orthography is an unstandardised but internally consistent adaptation of Russian cursive. In the letters we observe the following habits: the coda fricative is x (яхка *yahka* ‘even if’); /c/ is written both ц and ч (цеп *cep* ‘fish’, нуця ~ нуця *nuca* ‘Russian’); б appears for /p/ after nasals (камби ~ кампи *kampi* ‘letter’); both и and the pre-reform i serve for /i/; and — most striking — the devoiced /i/ of syllables between voiceless consonants is simply not written: цсе for *cise* ‘house’, цкин for *cikin* ‘if’, цнукара for *ci=nukara* ‘we read it’. This matches the strong devoicing of unaccented /i/ that Chiri describes for the dialect (Chiri 1942: 第1章). Sentoku also mixes in Arabic numerals and undigested Russian words (помощникъ старосты ‘assistant village head’), exactly as a bilingual letter-writer would.

- (2) сонко аноманде цкин нукараконде ва,

sonko an=omante cikin nukara konte wa
message INDEF.A=send if see give FIN

‘I send you this letter; when it arrives, please read it.’ — Tangiku & Ogihara 2001: Letter 1; East Sakhalin, Otasan (Sentoku Tarōji)

цкин with the devoiced i of cikin unwritten; nukara konte, literally ‘give seeing’, is a benefactive request.

- (3) Пороноп окта 4 цсе нуця ан

Poronop okta 4 cise nuca an
Poronop at four house Russian exist.SG

‘At Poronop there are four Russian households.’ — Tangiku & Ogihara 2001: Letter 1; East Sakhalin, Otasan (Sentoku Tarōji)

Sentoku writes okta where West-coast sources have ohta; both are attested forms of the locative postposition.

3.2 Latin

3.2.1 Pilsudski’s practical orthography The twenty-seven East-coast texts of Pilsudski’s *Materials* are printed in a practical Latin orthography which he explains in the prefatory “Remarks on Ainu phonetics, and on the transcription of the texts” (Pilsudski 1912: 1–10). Its values are Polish rather than English: *j* is the glide /y/, *v* alternates with *w*, and a palatalised sibilant “between *s* and *sh*” is written with the Polish acute, wavering with plain *s* in his own hearing — *isam* ~ *isam* ‘there is not’, *śumari* ‘fox’ (Pilsudski 1912: 5–6). The letter *x* renders the velar fricative that he correctly identified as found “only in Saghalien”, corresponding to Hokkaidō *p*, *t*, *k*: *axto* ‘rain’ beside Yeso *apto*, *cux* beside *cup* ‘sun’, *cex* beside *cep* ‘fish’ (Pilsudski 1912: 7, 9). Voiced letters *b*, *d*, *g* appear wherever he heard intervocalic voicing; stress is marked on all non-oxytone words; whispered word-final vowels are set as small superscript letters; and words are divided etymologically rather than by sandhi (Pilsudski 1912: 9–10). Dal Corso has recently mapped this notation systematically onto a phonemic analysis (Dal Corso 2024). Example (4) shows one of his sentences in the original orthography and in ours.

(4) *Ćiše oxta eširep án.*

<i>cise</i>	<i>oh̄ta</i>	<i>e-sirepa=an</i>
house	in	APPL-arrive=INDEF.S

‘I arrived at the house.’ — Pilsudski 1912: 160; Dal Corso 2024: 71; East Sakhalin

Pilsudski’s *ć* and *ś* record the automatic palatalisation of *c* and of *s* before *i*; the prefix *e-* on *sirepa* continues the deictic *he-* ‘up to’ (Dal Corso 2024: 71–72).

3.2.2 Linguistic romanisations Chiri Mashiho (知里真志保) gave the Sakhalin dialect its first systematic linguistic transcription. In his 1942 grammar he replaced the Hepburn-style digraphs of earlier work with *č*, *š*, *j* and *ü*, and wrote the Sakhalin coda fricative as *x* (*max* ‘woman’, *tex* ‘hand’) (Chiri 1942: 凡例); vowel length he left unwritten, regarding it as non-distinctive (Chiri 1942: 第 1 章), a decision since reversed by all later workers (see Itabashi (2001) on the phonemic status of length). Nakagawa traces how this notation fed into the strict one-letter-one-phoneme romanisation established in the 1950s and codified in the dialect dictionary edited by Hattori Shirō (服部四郎) (Nakagawa 2006: 3–5): there every vowel-initial syllable is preceded by an apostrophe for the glottal stop, /*c*/ is written *c*, the coda fricative *h*, and long vowels are doubled (Hattori 1964: 34).

Murasaki Kyōko (村崎恭子) adopted the Hattori system, in simplified form, for her grammar and for all her text editions: glottal stop ‘before vowel-initial syllables, doubled long vowels, particles written as separate words Murasaki (1976), Murasaki (2025: 凡例). Her transcription of the tales of Asai Take (浅井タケ) is the single largest body of Sakhalin text in this notation (Murasaki 2001). Recent philological work on the older corpora (Sakaguchi, Dal Corso) keeps the same letter values but omits the apostrophe and separates person clitics with an equals sign (Sakaguchi 2021: 45), which is also the practice of this grammar.

3.3 Katakana

Kana transcription of Ainu reaches back to the Edo-period interpreters, and native kana writing — Yamamoto Tasuke (山本多助), Nabesawa Motozō (鍋沢元蔵), Sunazawa Kura (砂沢クラ) and others — has been studied in detail by Kirikae (1997) and Nakagawa (2006: 24–31); that material is almost entirely Hokkaidō. For Sakhalin, kana enters with the Japanese administration. Kindaichi Kyōsuke’s (金田一京助) edition of the autobiography of the Sakhalin Ainu Yamabe Yasunosuke (山辺安之助), *あいぬ物語* (1913), prints the Ainu text as katakana ruby over the Japanese translation — though the spelling there is Kindaichi’s, not Yamabe’s own (Nakagawa 2006: 25). Sentoku, who had written Ainu in Cyrillic in 1906, used katakana for the Ainu names and words in his Japanese-language book of 1929 (Sentoku 1929).

The Sakhalin-specific kana conventions in use today were fixed by Chiri’s 1942 grammar: syllable codas are written with small kana (プ ツ ク シ ム for -p, -t, -k, -s, -m), and the coda fricative /h/ with a small *ha*-row kana agreeing in vowel colour with the preceding vowel — マハ *mah* ‘woman’, クフ *kuh* ‘belt’, テヘ *teh* ‘hand’, オホ *oh* ‘spear’ (Chiri 1942: 凡例). The vowel-matched small kana for coda *h* remain the standard way of writing Sakhalin words in kana and are used on this site (§3.5).

The main Sakhalin text actually written down in kana by an Ainu writer is the tale that Yamamoto Tasuke (1904–1993), a Kushiro Ainu and later a prolific kana author, recorded in February 1937 from Kimura Usarusima (木村ウサルシマ) of Ay on the East coast, and published in his Ainu-language journal アイヌ・モシリ (Sakaguchi 2019). Yamamoto writes codas with small kana, but uses シュ in words like クシュ for *kusu* ‘in order to’, and sometimes -p and -k (チカプ, パテク) where Sakhalin -h is expected — possibly carrying over his native Kushiro forms (Sakaguchi 2019: 114).

(5) オブケ マヌ。

<i>ohke</i>	<i>manu</i>
fart	REP

‘He farted, it is said.’ — Sakaguchi 2019: 115; East Sakhalin, Ay (Kimura Usarusima)

Yamamoto’s プ where the Sakhalin form is *ohke*; Sakaguchi retranscribes with -h.

(6) イコロ ヌカラ クシュ オマン。

ikoro *nukara* *kusu* *oman*
treasure see PURP go.SG

‘She went to look at the treasures.’ — Sakaguchi 2019: 116; East Sakhalin, Ay (Kimura Usarusima)

3.4 The romanisation of this grammar

This grammar writes Sakhalin Ainu with eleven consonant letters — *p t k c s h m n r w y* — and five vowel letters *a e i o u*, with long vowels doubled (*itah* ‘speech’, *kestoono* ‘the other day’). The glottal stop that begins every vowel-initial syllable is predictable and is left unwritten. The letter *h* serves both as syllable onset and for the coda fricative (*itah*, *ahkas* ‘walk’); other codas are written as the sources have them (*isam*, *kotan*; on the -p, -t, -k retained in the far south see Sakaguchi (2021: 45) and morphophonology). Person clitics are separated with an equals sign (*an=nukara*), affixes in morphemic lines with a hyphen. Table 1 maps these letters onto the source transcriptions discussed above.

This grammar	Hattori 1964 / Murasaki	Chiri 1942	Piłsudski 1912	Dobrotvorskij 1875 / Sentoku 1906
<i>p</i>	p	p	p, b	п, б
<i>t</i>	t	t	t, d	т, д
<i>k</i>	k	k	k, g	к, г
<i>c</i>	c	č	c, ć	ч, ц
<i>s</i>	s	s, š	s, ś	с, съ, ш
<i>h</i> (onset)	h	h	h	х, г
<i>h</i> (coda)	h	x	x	х
<i>m, n</i>	m, n	m, n	m, n	м, н
<i>r</i>	r	r	r (tr, dr after n)	p; initial тр
<i>w</i>	w	ü	w, v	в
<i>y</i>	y	j	j	й; я, ю, е for ya, yu, ye
<i>a e i o u</i>	a e i o u	a e i o u	a e i o u	а э/е и/і о у
long vowels <i>aa, ii...</i>	doubled	not written	stress accent only	stress accent only
glottal stop: unwritten	'	unwritten	unwritten	unwritten

Table 3: Table 1. The romanisation of this grammar and its source correspondences

Retranscription from each source therefore follows mechanically. From Murasaki-edited texts we drop the apostrophes and keep everything else, including the doubled vowels; from Piłsudski we replace *ś, ć, j, v, x* with *s, c, y, w, h*, remove the accents, and restore vowel length only where independent evidence supports it; from the Cyrillic sources we apply the table above, restoring the devoiced vowels that Sentoku leaves unwritten. Examples (1) and (4) above show the procedure for Cyrillic and for Piłsudski’s Latin; (7) completes the set with a sentence from Murasaki’s edition of Asai Take.

(7) *Sannupista re monimahpo 'okayahci.*

Sannupista *re* *monimahpo* *okay-ahci*
Sannupis.LOC three girl exist.PL-PL

‘In Sannupis there lived three girls.’ — Murasaki 2001: text 4; West Sakhalin, Odasu (Asai Take)

Murasaki’s apostrophe before ‘okayahci marks the glottal onset; we drop it. On the plural suffix -(a)hci see the chapter on number.

Where sources disagree on a form, the example keeps the form of the text actually cited, and the variation is discussed in prose (so Sentoku’s *okta* stands beside Dobrotvorskij’s *ohta* in §3.1). We never normalise one writer’s dialect to another’s. Restoring a phoneme the writer left unwritten — Sentoku’s devoiced *i* in *cikin* — is recovery *within* his own dialect, a different operation from rewriting one dialect’s form as another’s; the

first is phonemic restoration, the second is the normalisation we avoid. Plural verb morphology such as *-(a)hci* in (7) is treated in the number chapter, and the person clitics that the equals sign separates in person and alignment.

3.5 Script correspondences on this site

Outside the grammar, this site displays every Ainu form in three scripts — Latin, Cyrillic and katakana — converted automatically from the romanisation just described. The conversion between Latin and Cyrillic is lossless in both directions; the kana conversion is lossy, since broadly used Ainu katakana does not distinguish, for example, *ay* from *a.i* (both アイ). A form that cannot be syllabified is left unconverted. Table 2 gives the phoneme correspondences; Table 3 illustrates them with attested words.

Phoneme	Latin	Cyrillic	Kana, onset	Kana, coda
/p/	<i>p</i>	п	パピ プ ペ ポ	プ
/t/	<i>t</i>	т	タ テ ト; <i>tu</i> = ト ^o	ツ
/k/	<i>k</i>	к	カ キ ク ケ コ	ク
/c/	<i>c</i>	ц	チャ チ チュ チェ チョ	—
/s/	<i>s</i>	с	サ シ ス セ ソ	シ
/h/	<i>h</i>	х	ハ ヒ フ ヘ ホ	ハ ヒ フ ヘ ホ after a, i, u, e, o
/m/	<i>m</i>	м	マ ミ ム メ モ	ム
/n/	<i>n</i>	н	ナ ニ ヌ ネ ノ	ン
/r/	<i>r</i>	р	ラ リ ル レ ロ	ラ リ ル レ ロ after a, i, u, e, o
/w/	<i>w</i>	в	ワ ウィ ウェ ウォ	ウ
/y/	<i>y</i>	я, ю, ё, e for ya, yu, yo, ye; coda й	ヤ ュ ヨ イ エ	イ
/a i u e o/	<i>a i u e o</i>	а и у э о	アイ ウ エ オ	—
long vowels	<i>aa, ii...</i>	аа, ии...	ア ア (vowel kana repeated)	—

Table 4: Table 2. Latin–Cyrillic–katakana correspondences used on this site

Latin	Cyrillic	Katakana	Gloss
<i>itah</i>	итах	イタハ	speech
<i>aynu</i>	айну	アイヌ	person, Ainu
<i>cise</i>	цисэ	チセ	house
<i>ceh</i>	цэх	チェヘ	fish
<i>tuytah</i>	туйтах	トイタハ	folktale
<i>nispa</i>	ниспа	ニシパ	gentleman, sir
<i>yahka</i>	яхка	ヤハカ	even if
<i>ohta</i>	охта	オホタ	in, at

Table 5: Table 3. Attested words in the three scripts of this site

Two cells deserve comment. The kana coda series for /h/ and /r/ are vowel-matched, the convention introduced by Chiri (1942: 凡例) (§3.3); and the one-letter syllable ト^o for *tu* continues a long line of single-character devices for this syllable — ツ in Chiri's kana (Chiri 1942: 凡例), ト^o and ヲ^o in much twentieth-century practice, including Yamamoto Tasuke's (Nakagawa 2006: 26). The Cyrillic of the site is a phonemic transliteration of the

romanisation, not a revival of Dobrotvorsky's or Sentoku's spellings: it writes ə for /e/ in all positions except after /y/, keeps doubled long vowels, and uses no stress marks or hard signs.

Part II · Phonology

Chapter 4 · Phonology

Consonants and vowels, syllable structure, the coda system and h-neutralisation, vowel length, and prosody. Sakhalin Ainu has a compact segmental inventory — eleven consonants and five vowel qualities — combined with two properties that together give the variety its characteristic sound: a contrast of vowel length, and a strict limit on syllable codas that converts the stops **p *t *k* (and often **r*) into *h* at the end of a syllable. The first systematic description is Chiri (1942: §§1–20), written largely on the basis of East-coast speech; the West-coast dialects of Rayciska and Odasu were analysed phonologically by Hattori (Hattori 1967) and (Murasaki 1979: 1–10), and the East-coast dialects recorded by Pilsudski (1912: 1–9) in 1903–04 have been reassessed by Dal Corso (2024: 59–66). On vowel length and its history see Itabashi (2001); Nakagawa sets the same coda neutralisation and the systematic correspondence between Sakhalin vowel length and Hokkaidō accent in their pan-Ainu frame (Nakagawa 2024: 25, 45–49). Examples are given in the romanisation defined in Chapter 3; phonetic detail is added in square brackets where the sources supply it.

4.1 Consonants

	bilabial	alveolar	alveolo-palatal	velar	glottal
plosive	p	t		k	(ʔ)
affricate			c [tɕ]		
fricative		s			h
nasal	m	n			
tap		r [ɾ]			
approximant	w		y		

Table 6: Consonant phonemes of Sakhalin Ainu (after Chiri 1942: §2; Hattori 1967; Murasaki 1979; Dal Corso 2024: 59–64)

4.1.1 Obstruents and the absence of a voicing contrast Voicing is not contrastive anywhere in the system. Chiri (1942: §2(1)) states flatly that no pair of words is distinguished by voicing: the plosives and the affricate are voiceless in careful speech, but voiced variants [b d g dz] appear freely after nasals, as in his *membe* ‘grandmother’, *unji* for *unci* ‘fire’ and *henge* for *henke* ‘grandfather’. Pilsudski heard the same instability on the East coast: between vowels and after *y* the plosives “waver” between the voiceless and voiced series, which Dal Corso (2024: 61) renders as partially voiced [p̚ t̚ k̚] (e.g. *pájgara* for *paykara* ‘spring’). The two coasts are reported to differ in degree: Pilsudski found voicing commoner in the west, Chiri likewise heard more voiced tokens on the West coast, while in Murasaki’s West-coast recordings of the 1960s voicing of *p* is unattested and even *t k* are rarely voiced ((Dal Corso 2024: 61)). Dal Corso suggests contact with Japanese as one reason for the retreat of voicing in the later data.

The affricate *c* is alveolo-palatal [tɕ]: Pilsudski insists that it is neither [ts] nor English-type [tʃ] but “akin to the sound of the Polish *ć*” ((Pilsudski 1912: 7); (Dal Corso 2024: 63–64)). After nasals it has the voiced variant [dz], as in *únzi* for *unci* ‘fire’. A regular source of *c* is /t/ before *i*, where *t* shifts to *c* without exception, e.g. *kuci* ‘his belt’ from *kut-i* ((Chiri 1942: §16,6)); this feeds the coda alternations described in Chapter 5. In the East-coast dialects of Ay and Tunayci, *p* and *k* (but not *t*) are palatalised before *e* [ɛ], e.g. *emújke* for *emuyke* ‘all’, *tam̄be* for *tanpe* ‘this thing’ (Dal Corso 2024: 61–62); nothing comparable is recorded for the West coast.

The sibilant *s* varies between [s] and [ɕ]. Before and after *i*, and generally in coda position, it is [ɕ] on both coasts (*śine* for *sine* ‘one’ and *íso* for *iso* ‘bear’ next to *i*, and *ahkaś* for *ahkas* ‘walk’ in coda); elsewhere word-

initially the two sounds vary freely, Piłsudski recording both *su*j and *śu*j for *suy* ‘again’ (Piłsudski 1912: 5); (Dal Corso 2024: 62–63)). Chiri adds that plain [s] is the usual West-coast value, [ç] being obligatory only next to *i* and frequent in codas (Chiri 1942: §2(2)).

The glottal fricative *h* occurs both as onset and — its most distinctive role — as coda. As a coda it is a fricative whose place is determined by the preceding vowel: velar [x] after *a o e* (*mah* [max] ‘woman’), palatal [ç] after *i*, and labialised, near [ɸ], after *u* (*uh* [uɸ] ‘take’); Chiri, who writes this segment *x* and counts it a separate phoneme, notes that utterance-finally it is so weak as to escape notice Chiri (1942: §2(3)). We follow Hattori and Murasaki in treating onset [h] and the coda fricatives as one phoneme /h/. As onset before or after *u*, /h/ is [ɸ] systematically on the West coast but only variably in Piłsudski’s data (*fura* beside *húxkara*), and he also heard voiced [ɦ] between vowels ((Dal Corso 2024: 63)).

4.1.2 The tap *r* The single liquid is a tap [ɾ] between vowels. Word-initially, and after *n*, Piłsudski describes an articulation made “simultaneously” with a *t* placed further back than usual, pointing to a retroflex [ɽ] or affricated [tɽ], with much variation between speakers ((Piłsudski 1912: 8–9); (Dal Corso 2024: 64)). The hardening is lexicalised in part of the Sakhalin area: *tetara* for *retara* ‘be white’ is widespread, and the northern East-coast dialect of Nairo turns initial *r* into *t* in a whole set of words (*tup* for *ruh* ‘ice’) (Tangiku 2022: §2.6). In codas *r* does not occur: where Hokkaidō has syllable-final *r*, Sakhalin shows either an echo vowel or *h* (§4.4.4).

4.1.3 Nasals and approximants The nasals are *m* and *n*. Coda *m* survives only in careful speech: Chiri (1942: §16,2) observes that word-final *m* is normally pronounced *n* (*isan* for *isam* ‘not exist’, *aman* for *amam* ‘grain’), and that before *k t n* the shift is categorical (*enko* ‘half’, *kinta* ‘in the mountains’). Murasaki’s corpus bears this out: Asai Take’s tales have *isan teh an* beside slower *isam*. Conversely *n* assimilates to following labials and glides (*pompe* from *pon-pe* ‘little thing’; (Chiri 1942: §16,3)); the fuller sandhi picture is taken up in Chapter 5. Piłsudski additionally records palatal [ɲ] for /n/ before *e* and word-finally (*etunñe* ‘not want’), and a rare [ŋ] (Dal Corso 2024: 62).

The approximants are *w* and *y*. In the East-coast corpus /w/ is regularly written *v*, a labiodental value Piłsudski reports for all the dialects he documented; Dal Corso treats [v] as an East-coast allophone of /w/ rather than a separate phoneme (Dal Corso 2024: 63). Syllable-finally the two approximants form the falling diphthongs *ay uy ey oy* and *aw iw ew ow*; that these offglides count as consonants, not as vowel sequences, is shown by morphology — stems in *-y -w* select the consonant-stem allomorphs of suffixes, e.g. nominalising *-pe* rather than *-h* (*ray-pe* ‘dead one’), a point already made as a “morphophoneme” argument by Chiri (1942: §2(5)).

4.1.4 The glottal stop Every vowel-initial word begins phonetically with a glottal stop, strongest in accented syllables; Chiri regards it as fully predictable and therefore not a phoneme Chiri (1942: §1(3)). Hattori’s phonemicisation of Rayciska includes /ʔ/ in the consonant chart but bars it from codas (reported in (Tangiku 1998: §2-1)), and Dal Corso admits it for West Sakhalin only “provisionally, in free variation” (Dal Corso 2024: 59–60). Piłsudski never marks it, which may reflect his self-confessed “Slav ear” rather than its absence in the East (Dal Corso 2024: 60). We write no glottal stop: word-initially it is automatic, and word-internally its only role is to keep adjacent vowels apart at morpheme boundaries (see Chapter 5). Murasaki’s apostrophes (*’oman*, *’e’ani*) are accordingly not reproduced in our examples.

- (1) *Nea niśpa emus stómuśite.*

<i>nea</i>	<i>niśpa</i>	<i>emus</i>	<i>sitomusite</i>
that	gentleman	sword	gird.on

‘That gentleman girded on a sword.’ — Piłsudski 1912: 77; Dal Corso 2024: 67; East Sakhalin

Piłsudski’s *ś* shows the [ç] allophone of /s/ next to *i* and in coda position; his spelling *stómuśite* for *sitomusite* drops the initial unaccented *i*, which is devoiced between voiceless consonants (Chiri 1942: §1(4)).

4.2 Vowels

	front	central	back
high	i		u
mid	e		o
low		a	

Table 7: Vowel phonemes (each with a long counterpart: aa, ee, ii, oo, uu)

The five qualities *a e i o u* are constant across all Ainu varieties (Chiri 1942: §1). The back vowel *u* is only weakly and approaches [o], which is why early hearers wrote *ayno* for *aynu* and *inao* for *inaw* ‘ritual shaving’ (Chiri 1942: §1(1)). The mid vowels differ by coast: West Sakhalin *e o* are close-mid [e o], while Pilsudski describes open-mid [ɛ ɔ] for the East coast ((Pilsudski 1912: 1); (Dal Corso 2024: 60)). In unaccented syllables the mid and high vowels approach one another — Pilsudski’s *e o* “narrow” and his *i u* “open”, hence frequent interchange of *e~i* and *o~u* in the Ay and Tunayci texts and in Chiri’s material ((Chiri 1942: §19,1)). A more exotic East-coast allophone is [i] for /e/ in the causative *-te* after a voiced nasal-stop cluster, as in *čivéndy* for *ciwen-te* ‘was destroyed’, a sound Pilsudski compares to Russian ы (Dal Corso 2024: 60). Unaccented *i* between voiceless consonants devoices, in Sakhalin conspicuously so (Chiri 1942: §1(4)). Finally, where Hokkaidō has word-final *-e* several Sakhalin words show *-i*: *tani* ‘now’, *poni* ‘bone’ (Hokkaidō *tane*, *pone*) (Tangiku 2022: §2.1).

4.3 Vowel length

Length is the hallmark of Sakhalin Ainu vocalism. Hattori established for Rayciska that long and short vowels contrast ((Hattori 1967)), and Murasaki’s grammar and text collections write the contrast consistently ((Murasaki 1979); (Murasaki 1976)). Chiri, working mainly from East-coast speech, had denied any contrast, treating long vowels as a by-product of accent Chiri (1942: §1(2)); Tangiku (1998: §2-3) attributes the disagreement to the difference between the two dialect groups, the East-coast prosody never having been adequately recorded. For the West coast at least the contrast is real, though its functional load is modest; some pairs follow.

short		long	
<i>ku=</i>	‘I, my’	<i>kuu</i>	‘bow’; ‘to drink’
<i>sine</i>	‘one’	<i>siine</i>	‘to rest’
<i>nisahta</i>	‘in the morning’	<i>niisahto</i>	‘suddenly’
<i>hekaci</i>	‘boy, child’	<i>heekopo</i>	‘younger sister’

Table 8: Short–long (near-)minimal pairs (Hattori 1964, as collated in Itabashi 2001: 89, 96, 101)

Two phonetic regularities interact with the contrast. Monosyllabic content words of the shape (C)V are always long — *nii* ‘tree’, *kuu* ‘bow’, *ree* ‘name’, *too* ‘day, lake’, *poo* ‘child’ — whereas clitics and affixes of the same shape are short ((Chiri 1942: §1(2)); (Itabashi 2001: 89)). And some of these monosyllables shorten under suffixation, e.g. *kii* ‘do’ but *ki-hci* ‘they do’, which leads Tangiku (2022: §2.1) to analyse the vowel of such roots as phonemically short with surface lengthening; the long vowels of polysyllables (*heese* ‘breathe’, *niina* ‘gather firewood’) do not alternate and must be underlying.

- (2) *nee* *kusu* *neyke* *unciuwaare* *aynu* *unciuwaare*
 COP because then fire.kindle person fire.kindle

‘and so the fire-tender tended the fire’ — Murasaki 2001: text 18; West Sakhalin, Odasu (Asai Take)

The tale continues *niina aynu niina* ‘the firewood-gatherer gathered firewood’: the long vowels of *uwaare* ‘kindle’ and *niina* stand against the short vowels of *aynu* and *unci*.

Historically, Sakhalin length corresponds in large measure to the pitch accent of the Hokkaidō dialects: where a Hokkaidō word has an irregular high accent on an open first syllable, the Sakhalin cognate has a long vowel there — *ká* : *kaa* ‘thread’, *hése* : *heese* ‘breathe’, *réra* : *reera* ‘wind’, *túnas* : *tuunas* ‘be early’ ((Itabashi 2001: 93–95); (Hattori 1964)). On this basis Hattori reconstructed Proto-Ainu with distinctive length and no accent, Sakhalin being conservative (Hattori 1967). The correspondence is not one-to-one, however: Itabashi (2001: 95) lists pairs such as Hokkaidō *niná* ‘crush’ (accent on the second syllable) against Sakhalin *niina*, or *nociw* : *noociw* ‘star’, and concludes — with Vovin (1993) — that the proto-language needed both length and pitch as

independent features. The opposite position, that Sakhalin length is a secondary development from accent, is argued by Satō (Satō 2015), who reconstructs Proto-Ainu with pitch accent and no length, on the strength of the Japanese loan *tuuki* ‘lacquered cup’ (long *uu* matching the irregular accent of Hokkaidō *túki*, though the Japanese source vowel is short) and of alternations like *nii* ‘tree’ beside the compound *ah-ni* ‘elm’ (Bugaeva 2022). The debate does not affect the synchronic statement: in the recorded West-coast language, length distinguishes morphemes and must be written.

4.4 Syllable structure and phonotactics

The syllable is (C)V(V)(C). Onsets contain at most one consonant — there are no clusters and no onset glide sequences of the Japanese ㄱㅇ type — and an apparently onsetless syllable receives a glottal onset ((Chiri 1942: §§3–5); (Itabashi 2001: 89)). The nucleus is a short or long vowel; long vowels occur freely in open syllables but a coda forces a short nucleus, so that possessive suffixation shortens a long stem vowel (*imii* → *imi-yehe* ‘garment’, Chapter 5). Word-internal consonant sequences are exhausted by coda + onset: *ahkas* ‘walk’, *tehkuh* ‘wing’, *muysankeh* ‘broom’. Sequences of unlike vowels within a morpheme are rare, and at morpheme boundaries hiatus is repaired by glide or *h* insertion (Chapter 5).

4.4.1 Syllable types and the bimoraic minimum Five syllable shapes occur: V, CV, VC, CVC, and — the shape that sets Sakhalin apart from Hokkaidō — CVV with a long nucleus ((Itabashi 2001: 89)). The two vowel-initial types are restricted: in native simplex words V and VC syllables stand only word-initially, and where word formation would create one elsewhere it is absorbed into the preceding syllable (*sik-o* ‘eye attaches’ > *siko* ‘be born’, *rup-us* ‘ice attaches’ > *rupus* ‘freeze’) or removed by haplology or vowel deletion (*kera-an* > *keran* ‘be tasty’). Chiri accordingly reduces the basic templates to CV and CVC, with the onsetless types as word-initial variants carrying a zero onset (Chiri 1942: §5).

type	example	gloss
V	<i>o·ro</i>	place
CV	<i>ko·tan</i>	village
VC	<i>ah·kas</i>	walk
CVC	<i>sik</i>	eye
CVV	<i>nii</i>	tree

Table 9: The five syllable types, with the relevant syllable shown first (a raised dot marks the syllable boundary); V and VC occur only word-initially

Following Murasaki’s moraic description, Itabashi (2001: 89) treats Sakhalin as a mora-counting language: a CV sequence, a bare V, and a coda consonant each contribute one mora, and the minimal accent-bearing free form is two moras. This bimoraic minimum is what lies behind the lengthening of open monosyllables noted in §4.3: a content word of the bare shape (C)V would be a single mora, too light to carry the accent it requires, and so surfaces as CVV — *nii* ‘tree’, *kuu* ‘bow’, *too* ‘day, lake’ — while clitics and affixes, which are unaccented, escape the requirement and stay short. The same arithmetic excludes long vowels from closed syllables: CVVC would run to three moras, and the language repairs it by shortening (*imii* → *imi-yehe* above).

4.4.2 The coda inventory The coda is where Sakhalin phonotactics departs most sharply from Hokkaidō. Hokkaidō Ainu closes syllables with any of nine consonants, *p t k s r m n w y*; Sakhalin permits only six, *s n m w y h*, with coda *m* marginal (§4.1.3) — the plosives, *c*, *r* and the glottal stop are excluded (Hattori 1967, as reported in (Tangiku 1998: §2-1); (Itabashi 2001: 89)). The northern East-coast dialects from Taraika to Nairo, together with Tarantomari on the West coast, stand outside the restriction and keep Hokkaidō-style *p t k* (Tangiku 2022: §2.3); they are in this respect the most conservative Sakhalin varieties (§4.4.3).

variety	permitted codas	excluded codas
Hokkaidō	p t k s r m n w y	c h ?
Sakhalin (general)	s m n w y h	p t k c r ?
Taraika–Nairo; Tarantomari	p t k s ... (as Hokkaidō)	—

Table 10: Coda inventories (Chiri 1942: §6; Hattori 1967, apud Tangiku 1998: §2-1; Tangiku 2022: §2.3)

Chiri drew the quantitative consequence of this restriction in a small piece of arithmetic Chiri (1942: §6). The vowels supply 5 onsetless V syllables; the 11 onset consonants combine with the 5 vowels into 51 CV syllables ($11 \times 5 = 55$, minus the four non-occurring *ti wi wu yi*); these figures are common to both areas. The codas then multiply out differently: VC gives $5 \times 9 = 45$ syllables in Hokkaidō but $5 \times 6 = 30$ in Sakhalin, and CVC gives $51 \times 9 = 459$ against $51 \times 6 = 306$. In total Hokkaidō has 560 possible basic syllables and Sakhalin 392 — the coda restriction deletes nearly a third of the syllable space. (Chiri counts his Sakhalin-only coda fricative *x*, our *h*, among the six; and his count has no CVV type because he treated length as non-distinctive, §4.3. Adding the long-vowel syllables would not affect the comparison, which turns entirely on the codas.)

type	Hokkaidō	Sakhalin
V	5	5
CV	51 ($11 \times 5 - 4$)	51
VC	45 (5×9)	30 (5×6)
CVC	459 (51×9)	306 (51×6)
total	560	392

Table 11: The basic syllable inventory (Chiri 1942: §6)

4.4.3 Coda neutralisation The signature phonotactic rule of Sakhalin Ainu is the neutralisation of coda obstruents. Syllable-final **p *t *k* all surface as *h* (phonetically [x ɕ φ] by vowel environment, §4.1.1), except that the coronal and dorsal **t *k* after *i* palatalise to *s* (*sik* → *sis* ‘eye’); labial **p* gives *h* even after *i* (*cip* → West *cih* ‘boat’, East *cis*) ((Chiri 1942: §16,1); (Tangiku 2022: §2.3)). Final **r* goes its own way and is treated in §4.4.4. The rule applies word-internally as well as finally (*ahkas* ‘walk’ : Hokkaidō *apkas*; *yuhke* ‘be strong’ with root *yup*, (Dal Corso 2024: 62)), and it is the engine of the stem alternations treated in Chapter 5, since the underlying stop reappears before vowel-initial suffixes.

A second source of coda *h* is the resolution of geminates and other stop clusters. Where Hokkaidō has a long stop or a heterorganic stop sequence, Sakhalin has *h* plus the following stop: the geminates *wakka* : *wahka* ‘water’, *satte* : *sahte* ‘dry’, and the cluster *hotke* : *hohke* ‘lie down’ ((Itabashi 2001: 93–94); (Chiri 1942: §16,8)). Nothing new needs to be stated: the first stop has no following vowel of its own, so it debuccalises by the rule just given, while the second survives as the onset of the next syllable. Geminates in loanwords are reshaped the same way (§4.4.5).

Hokkaidō	Sakhalin	
<i>itak</i>	<i>itah</i>	‘speech; speak’
<i>cikap</i>	<i>cikah</i>	‘bird’
<i>mat</i>	<i>mah</i>	‘woman, wife’
<i>cup</i>	<i>cuh</i>	‘sun, moon’
<i>sik</i>	<i>sis</i>	‘eye’
<i>cip</i>	<i>cih</i> (E <i>cis</i>)	‘boat’
<i>kisar</i>	<i>kisara</i> ~ <i>kisaru</i>	‘ear’
<i>mukar</i>	<i>mukara</i>	‘axe’
<i>utar</i>	<i>utara</i> ~ <i>utah</i>	‘people’ (§4.4.4)
<i>unarpe</i>	<i>unahpe</i>	‘aunt’

Table 12: Coda correspondences, Hokkaidō : Sakhalin (Hattori 1964; Itabashi 2001: 93–94; Tangiku 1998: §2-2; Tangiku 2022: §§2.3–2.4)

The northern East-coast dialects from Taraika to Nairo, and Tarantomari on the West coast, stand outside the rule and retain *tek*, *sik*, *rup* like Hokkaidō (Tangiku 2022: §2.3). In Piłsudski's other East-coast dialects the final fricative was evidently faint: he usually writes nothing at all (*pate* for *pateh* ‘only’) and admits confusing *ee* ‘eat’ with *eh* ‘come’ and *ruu* ‘road’ with *ruh* ‘sea ice’, an error Chiri diagnoses precisely as the weakness of utterance-final [x] ((Chiri 1942: §16,1); (Dal Corso 2024: 61–62)).

- (3) *rihpo ka ahkascikah kayki taa tehku^{pi}yays^{ke} taa muys^{an}ke^{ma} sahkane naa koro- nee manu*
hi hci-
 height on walk bird even so wing- shed and so broom as also chop- as also ^{pe}have- COP REP
 POSS stick PL-NMLZ

‘even the birds that walk up above shed their wing-feathers, and they had them for brooms and for chopsticks, it is said’ — Murasaki 2001: text 19; West Sakhalin, Odasu (Asai Take)

A closing formula of Asai Take's tales. Neutralised codas throughout: *ahkas* (Hokkaidō *apkas*), *cikah* (*cikap*), *tehku*, *muysankeh*, *sahka*; *koro-hci* shows the open -rV stem of ‘have’.

4.4.4 Coda *r*: echo vowel and debuccalisation Coda *r* patterns apart from the stops. Chiri gives two resolutions for syllable-final **r*: either the preceding vowel is copied, opening the syllable — *utara* ‘people’, *kisara* ‘ear’, *mukara* ‘axe’, *koro* ‘have’ against Hokkaidō *utar*, *kisar*, *mukar*, *kor* — or the *r* debuccalises like a stop, word-finally in *utah* and before a consonant in *unahpe* ‘aunt’ (Hokkaidō *unarpe*) (Chiri 1942: §16,4). He presents the routes as alternatives, citing *utara* and *utax* from the very same page of Piłsudski's *Materials* — so no purely segmental conditioning can decide between them.

The corpus shows the choice to be in large part dialectal. In the East-coast texts the echo vowel is all but categorical. In the tale Piłsudski took down at Takoye from the nineteen-year-old Kusurikoya, *utara* recurs from one end to the other — *onneru utara* ‘the old people’, *mahnekuh utara* ‘the women’, *ohkayo utara* ‘the men’ — and debuccalised *utah* never appears; Sakaguchi, surveying the whole East-coast folktale corpus, states the generalisation directly: the variants *uta* ~ *utah* are West-coast forms, his eastern materials showing almost exclusively *utara* ((Sakaguchi 2020a: 187); (Sakaguchi 2020c)). The *h*-forms cluster in the West-coast records — Murasaki's conversational material and the modern revitalisation register that continues that norm. We therefore read the second route as a West-coast extension of debuccalisation from the obstruents to *r*: having converted **p *t *k* to *h*, the western varieties drew the one remaining eligible coda into the same lenition, while the East keeps *r* distinct by epenthesis. Even in the west the two forms coexist word by word (*utara* and *utah* are both current), so the residue is lexical variation rather than rule.

- (4) нэр хэка́ци утара́ вэ́ яй́ цисе́ коро, ну́рика ока́й ки-си.

neroh hekaci utara wee-yay-cise-koro pirika okay kii-si
 those child PL RECP.APPL-REFL-house-have good living do-PL

‘Those children each kept their own house and lived well.’ — Sakaguchi 2020c: 63; East Sakhalin, Takoye (Kusurikoya, rec. Pilsudski)

utara ‘people; PL’ with the echo vowel, the East-coast norm; the tale, recorded in Cyrillic, shows utara throughout and no utah.

- (5) *anpene neeraan hecire anahka anpene anoka pateh an=tek-ih ihi utah ninpa*
 trulyly whatever dance even trulyly me only INDEF.A=hand-POSS people grab

‘Truly, whatever the dance, it was always just me whose hand people grabbed.’ — Murasaki 1976: 82; Sakaguchi 2020a: 193; West Sakhalin, Maoka (conversation)

utah ‘people’ with the debuccalised coda, the West-coast reflex of utar; beside it Maoka speech also has apocopated uta (sake naa uta i-kuure ‘people made me drink sake too’, Murasaki 1976: 77).

The echo vowel is phonologically real, not an orthographic convenience: stems that acquired it (*koro* ‘have’, *nukara* ‘see’, Hokkaidō *kor*, *nukar*) still behave as consonant-final under suffixation, the hidden *r* resurfacing or assimilating (*koro* + *-te* → *konte* ‘let have, give’), in contrast with stems whose final vowel is original, such as *tura* ‘be with’ (Tangiku 1998: §§2–3). These alternations belong to the stem morphophonology and are taken up in Chapter 5.

4.4.5 Loanword adaptation Borrowings are reshaped to this template. Initial *r* and foreign *l* are avoided: Tungusic *luca* ‘Russian’ was borrowed as *nuca* (Chiri 1942: §17,2), and 蠟燭 ‘candle’ (Chinese, via Japanese *rōsoku*), through the intermediate *ratcaku*, appears as *cahraku* ‘lamp’ with metathesis of the initial *r* ((Chiri 1942: §20); Chapter 25 (Language contact and loanwords)). Japanese accent and vowel quantity are mapped onto length: *kane* ‘metal’ gives *kaani*, where Hokkaidō has accented *káni* (Tangiku 2022: §2.1), and *tuki* ‘cup’ gives *tuuki* (Bugueva 2022). Geminate *pp* surfacing as *hp* (Chapter 25). Contact vocabulary also flowed from the Amur–Sakhalin languages, e.g. *tanku* ‘hundred’ from Tungusic (Tangiku 2022: §3.3). Russian, by contrast, left almost no integrated lexicon: its mark on Sentoku Tarōji’s Cyrillic letters is chiefly the written apparatus of dates and sums and the personal and place names adapted to Ainu phonology — *Aleksandr* → *Ariksandoru* (with the cluster broken by an echo vowel and *l* → *r*), *lesnichiy* ‘forester’ → *Lesneciy* — rather than borrowed words (Chapter 25) (Tangiku & Ogihara 2001).

4.5 Prosody

Sakhalin Ainu has no distinctive pitch accent. Hattori and Murasaki agree that the West-coast dialects oppose no word pairs by pitch; what one hears is a predictable prominence pattern — the first syllable is high if it is heavy (a long vowel or a closed syllable), otherwise the second syllable is high: *héekopo* ‘younger sister’, *káhkemah* ‘lady’, *cikáh* ‘bird’, *sapá* ‘head’ ((Hattori 1967); (Murasaki 1979: 5–6); (Tangiku 2022: §2.5)). Murasaki describes the system in moraic terms — main prominence on the second mora, with secondary beats on later even-numbered moras, shifted leftward when the expected mora is the second half of a long vowel or a coda (as summarised in (Itabashi 2001: 90)). The contrastive work done by pitch in Hokkaidō is thus carried in Sakhalin by vowel length (§4.3). Chiri’s description, by contrast, sets up two lexical accent classes (peak on the first or on the second syllable) and distinguishes derivation from compounding by accent placement Chiri (1942: §§9–12); Tangiku (1998: §2-3) takes this to reflect his East-coast material, whose prosody was never instrumentally recorded. We treat the question of an East-coast accent contrast as open; the Hokkaidō correspondences (§4.3) predict that if any dialect kept pitch distinctive it would show up precisely in pairs like *niina* ‘crush’ ~ ‘gather firewood’, and Pilsudski’s accent marks, which Dal Corso leaves uninterpreted, are the only place to test it.

(6) *Tan kotàn ohtà utása=an kusú áriki=an.*

tan kotan ohta utasa=an kusu ariki=an
 this village in visit=INDEF.S PURP come.PL=INDEF.S

‘I have come to call on this village.’ — Dobrotvorskij 1875; Sakaguchi 2021; West Sakhalin, West coast, central (recorded 1867–1872)

Dobrotvorskij marked prominence only on words he recorded himself. *utása* carries the default second-syllable peak; *áriki* (from **arki*, with epenthetic *i*, cf. Chiri 1942: §18,2) is marked on the initial — we read this as the prominence of the originally closed first syllable **ar-*, the epenthetic vowel being invisible to accent assignment.

Phrase-level effects complete the picture: accented open monosyllables stretch (*ta-ta* ‘there-at’ pronounced *taata*, (Chiri 1942: §1(2))), final *h* fades before pause (§4.4.3), and emphasis lengthens vowels beyond their phonemic value (*poronno* ~ *poroono* ‘plenty’, (Chiri 1942: §18,3)).

4.5.1 The glottal onset as a prosodic signal The glottal stop of §4.1.4 belongs in part to prosody. Chiri’s formulation is prosodic from the outset — the onset appears on the initial vowel of an *accented* syllable, and is “especially conspicuous” in Sakhalin Chiri (1942: §1(3)) — and the West-coast orthographic tradition tracks the same conditioning: Murasaki’s apostrophes sit precisely on accent-bearing vowel-initial morphemes (*’itah*, *’oman*, and word-internally at junctures such as *’e’ani*), not on every orthographic vowel. For the East coast, where Pilsudski wrote nothing, the segment betrays itself indirectly, through sandhi. A word-final stop debuccalises only when no vowel follows; before a genuinely vowel-initial word it should resyllabify as an onset and keep its stop value. At the juncture *nah an* ‘being so’, however, Pilsudski hears [nax án]: the coda of *nah* is treated as pre-consonantal, which means the accented *an* must begin with a consonant — the glottal onset (Dal Corso 2024: 61). The accent-linked glottal stop is thus recoverable for the East from its phonological effect even though no transcriber recorded it directly.

(7) *nax án*

nah an
 so exist

‘to be so’ — Dal Corso 2024: 61; East Sakhalin

The coda of *nah* surfaces as the fricative [x] before *an* because *an* begins with a consonantal glottal onset, blocking resyllabification; the acute is Pilsudski’s accent mark.

4.5.2 Pitch, length, and Proto-Ainu Set against Hokkaidō, the Sakhalin system looks like one half of an old inheritance. The correspondences of §4.3 — Hokkaidō pitch matching Sakhalin length in the regular cases, with residues like *niná* : *niina* ‘crush’ that no single-feature account derives — point, with (Vovin 1993) and (Itabashi 2001: 95–98), to a proto-language in which vowel length and pitch were independent distinctive features (against Hattori’s length-only reconstruction, §4.3). Each daughter branch then kept one feature and made the other predictable. Hokkaidō phonemicised pitch, length surviving only as a phonetic correlate of the accented open syllable; Sakhalin phonemicised length, pitch surviving only as the automatic second-mora contour described above. The two systems carry the same old prominence by different means, which is why the cognate sets translate so regularly between an accent mark on one side of the strait and a doubled vowel on the other.

Kuril Ainu can be placed in this scheme only tentatively. Its record is confined to early wordlists — Krasheninnikov, Voznesenskii, Dybowski, Torii — which are fragmentary and mutually inconsistent on quantity: ‘two’ appears as *tuup* ~ *tuuh* ~ *tubic* across collectors. From this material Itabashi (2001: 90–91) nonetheless argues that Kuril had a pitch accent of the Hokkaidō type, a distinctive low-to-high rise, while the irregular long-vowel spellings leave open whether length was also contrastive. If that is right, the length-for-pitch reanalysis is a specifically Sakhalin innovation, not a shared northern development — and the pieces described in this chapter, the length contrast (§4.3), the bimoraic minimum (§4.4.1) and the accent-conditioned glottal onset, form a single prosodic complex organised around quantity.

Chapter 5 · Morphophonology

Alternations at morpheme boundaries: coda alternations, vowel-initial allomorphy, possessive-suffix vowel copying, and sandhi.

The coda restrictions of Chapter 4 make Sakhalin Ainu morphology audibly alternating: a stem that ends in *h* in isolation recovers its underlying stop before a vowel-initial suffix, a stem that ends in an echo-vowel -*rV* sequence behaves as if it ended in *r*, and the possessive suffix copies its vowel from the stem. This chapter describes these alternations and the sandhi of connected speech. The classic statement of the synchronic rules is the chapter on sound change in Chiri (1942: §§13–16); for the West coast the system is worked out by Murasaki (1979) and, for the verbs in -*rV*, by Tangiku (1998); the East-coast counterparts are surveyed by Dal Corso (2024: 65–66). A useful summary of the whole complex is Tangiku (2022: §§2.3–2.4).

5.1 Coda alternations: *h* with *p*, *t*, *k*

Stems whose final stop has been neutralised to *h* (Chapter 4) show the stop whenever a vowel follows within the word. The alternation runs through both nominal and verbal morphology. With the possessive suffix (Chapter 7): *teh* ‘hand’ but *tek-ih* ‘his hand’, *otoh* ‘hair’ but *otop-ih* (Tangiku 2022: §2.3). With person suffixes: *itah* ‘speak’ but *itak-an* ‘we spoke’ and *itak-ahci* ‘they spoke’ (Tangiku 1998: §§2-2, 6). Before *i* the recovered *t* is regularly affricated (Chapter 4): *mah* ‘wife’ : *mac-ih* ‘his wife’, *kuh* ‘belt’ : *kuc-i* ‘his belt’ (Chiri 1942: §16,6). And where the coda has shifted to *s* after *i*, the velar returns intact: *sis* ‘eye’ : *sik-ih* ‘her eyes’, the latter frequent in Asai Take’s tales (*sikihi masaha* ‘her eyes opened’, (Murasaki 2001: text 7)). We therefore set up underlying stem-final /p t k/ and let the neutralisation rule of Chapter 4 apply in coda position, in line with Chiri’s and Tangiku’s analyses and with Dal Corso’s notation *itak** for surface *itah* (Dal Corso 2024: 61).

citation	possessed		source
<i>teh</i>	<i>tekihi</i>	‘hand’	corpus; Tangiku 2022
<i>otoh</i>	<i>otopihi</i>	‘hair’	Tangiku 2022
<i>sewreh</i>	<i>sewrepihi</i>	‘windpipe’	Hattori 1964
<i>sistoh</i>	<i>sistokihi</i>	‘elbow’	Hattori 1964
<i>nisah</i>	<i>nisapuhu</i>	‘shin’	Hattori 1964; corpus
<i>mah</i>	<i>macihi</i>	‘wife’	corpus
<i>kosmah</i>	<i>kosmacihi</i>	‘daughter-in-law’	Hattori 1964
<i>sis</i>	<i>sikihi</i>	‘eye’	corpus

Table 13: Citation form and prevocalic (possessed) form of *h*-final nouns

- (1) *tani taa mukarani humpneampmukarkotahnowa tani taa teh ani sitaykineamppek- kotahma*
 now so axe with *ha* chop- then axe stick then now so hand with *hi* strike- then *ih* hand- stick
 NMLZ NMLZ POSS

‘now when she chopped it with the axe, the axe stuck; then when she struck it with her hand, her hand stuck’ — Murasaki 2001: text 43; West Sakhalin, Odasu (Asai Take)

Citation form *teh* before the vowel-initial word *ani*, possessed form *tek-ih* before the vowel-initial suffix: the stop is recovered word-internally only.

- (2) *orowa taa toy- tek- ehkempneampenisapuanise kayteh teh wahka onne ahun teh ray*
 aynu ihi ha apu-
 then so Earth.Mahand- pull-NMLZ then Spindle- shin- break and water into enter.SG and die.SG
 POSS POSS

‘then, when he pulled Earth-Man’s hand, Spindleshanks broke his shin, fell into the water and died’ — Murasaki 2001: text 10; West Sakhalin, Odasu (Asai Take)

nisapu-hu beside citation *nisah* ‘shin’ (Hattori 1964 *nisah*, -*puhu*): the recovered *p* brings its own lexical vowel *u*.

The same alternation underlies the basic numerals: *sine-h* ‘one (thing)’, *tu-h* ‘two’, *re-h* ‘three’ contain the nominaliser -*h*, the Sakhalin outcome of *-*p* (Hokkaidō *sinep*, *tup*, *rep*); Chiri already used the suffix pair -*pe* (post-consonantal) ~ -*h* (post-vocalic) as a diagnostic of stem shape (Chiri 1942: §2(5)) (see Chapter 10).

Across word boundaries usage differs by source. Chiri describes liaison: the stop surfaces before a vowel-initial word, as in *tek ani* ‘with the hand’, *ek-an* ‘I came’, *mat-ah* ‘younger sister’ (Chiri 1942: §16,1), and the letters of Sentoku Tarōji show spellings like *cep ukoyki* ‘fish-catching’ (Tangiku & Ogihara 2001: Letter 1). In Murasaki’s West-coast transcriptions, by contrast, the neutralised form stands even before a vowel — *teh ani* in the example above — which we attribute to the automatic glottal onset of the following word: liaison applies only where no glottal stop intervenes, hence within the word but not across it in these dialects. Between identical vowels and before stops the debuccalised segment may also surface between words in Piłsudski’s texts (*nax an* ‘being so’, *itax ki* ‘make a speech’) (Dal Corso 2024: 62).

5.2 The possessive suffix and vowel copying

The possessed (“belonging”) form of the noun is built with a suffix of the shape *-hV* on vowel-final stems and *-VhV* on consonant-final stems (Chapter 7 treats its syntax). On vowel-final stems the suffix vowel copies the stem-final vowel: *sapa* ‘head’ → *sapa-ha*, *cise* ‘house’ → *cise-he*, *mahpoo* ‘daughter’ → *mahpoo-ho* (Murasaki 1979); (Tangiku 2022: §4.1). On consonant-final stems the two vowels are identical to each other but lexically determined by the stem, with strong tendencies set by the final consonant — *e* after *y w*, *i* or *u* after nasals and stops: *tek-ih*, *ram-u*, *kam-ih* ‘meat’, *kotan-u* ‘village’ ((Murasaki 1979), as tabulated in (Tangiku 2022: §4.1); (Hattori 1964)). Sakhalin has only the long form: Hokkaidō’s short possessive (*tek-e*) has no West-coast counterpart (Tangiku 2022: §4.1).

Lengthened monosyllabic stems regularly shorten before *-he*, since the suffixed word is already bimoraic and the lengthening that made the free form heavy is no longer needed: *kuu* ‘bow’ ~ *ku-he*, *nii* ‘tree’ ~ *ni-he* (Chiri (1952: 463), (Chiri 1942: §72); Chapter 7). The shortening is not exceptionless, however: *ree* ‘name’ surfaces unshortened in the Yamada Hayo corpus (*reehe tah nee manu* ‘his name was this’, (Kitahara et al. 2003)). Polysyllabic stems in *ii uu* shorten the vowel and surface a glide before *-e*: *imii* ‘garment’ → *imi-yehe*, and *ruu* ‘road’ → *ruwe-he* (Tangiku 2022: §4.1), the coda ban on long nuclei (Chapter 4) forcing the shortening.

- (3) *kotanu-hu ruwe-he ohta oman teh tani taa ruu okakara oman ike taa ciseta oman*
village- road-POSS in go.SG and now so road along go.SG and so house.in go.SG
POSS

‘she went on the road to her village, went along the road, and came to the house’ — Murasaki 2001: text 25; West Sakhalin, Odasu (Asai Take)

Citation form *ruu* beside possessed *ruwe-he* in one sentence; *kotan* takes the lexical vowel *u*.

Two analyses of the suffix have been proposed. Chiri derives it historically from emphatic doubling of the final vowel with *h* inserted to break the hiatus: *ona* ‘father’ → *onaa* → *onaha*, *siki* → *sikii* → *sikihi* (Chiri 1942: §18,2). Dal Corso, for the East coast, takes the underlying form to be *-(i)hi*, with the first vowel of *-VhV* an optional epenthesis resolving the consonant + *h* cluster — supported by free variation of *kotanu* and *kotanh* in Piłsudski’s corpus — and with *h* deletable after nasals, giving *kotanu* (Dal Corso 2024: 66). The copy pattern of the West-coast data is stated most simply by Chiri’s account, which we adopt for this grammar while noting that the East-coast variation favours Dal Corso’s.

5.3 Person prefixes and clitics before vowel-initial stems

The person markers (Chapter 13) end in a vowel (*ku= e= eci= i=*) or in *n* (*an=*), and their combination with vowel-initial stems is the main laboratory of hiatus resolution. Vowel-final markers simply stand before the stem vowel, which keeps its glottal onset: Chiri transcribes *an-e* ‘we eat it’ as [ane:] and *e-e* ‘you eat it’ as [ee:], with the clitic vowel short and the accented stem vowel lengthened (Chiri 1942: §1(2)); Murasaki’s apostrophe spellings of the type *e’ani* ‘you’ make the same point for the 1960s data. In the nineteenth-century West-coast sentences recorded by Dobrotvorsky and in Sentoku Tarōji’s letters the pattern is directly attested with *an=* before vowel-initial stems.

- (4) *ka ponno ponno cep ukoyki an=e=hci*
even little little fish catch INDEF.A=eat=PL

‘we catch and eat only a very little fish’ — Tangiku & Ogihara 2001: Letter 1; East Sakhalin, Otasan (Sentoku Tarōji)
an= stands directly before the vowel-initial stem *e* ‘eat’; no glide or elision applies.

- (5) *Tan kotán ta eh renkáyne, an=è=núkara.*

tan kotan ta eh renkayne an=e=nukara
this village in come.SG because INDEF.A=2SG.O=see

‘now that you have come to this village, I meet you’ — Dobrotvorsky 1875; Sakaguchi 2021; West Sakhalin, West coast, central (recorded 1867–1872)

Two vowel-initial morphemes in sequence after *an=*; note also word-final *eh* ‘come’ (Hokkaidō ek).

- (6) *Eiráki.*

e=i=ray-ki
2SG.A=INDEF.O=die-CAUS

‘you are killing me’ — Pilsudski 1912: 125; Dal Corso 2024: 68; East Sakhalin

The 2SG marker *e=* before vowel-initial *i=*; in Pilsudski’s connected transcription the three vowels run together.

Identical vowels at a marker boundary may coalesce: in Asai Take’s tales *eci=itak-ahci* ‘they call you’ is realised *ecitakahci* ((Murasaki 2001: text 14)). We analyse this as the same-vowel contraction that Chiri documents inside compounds (*keran* ‘tasty’ from *kera-an* ‘flavour-exist’, *sinan* from *sine-an* ‘a certain’; (Chiri 1942: §14,1)). Unlike vowels in derivation are instead separated by a glide that copies the frontness or rounding of the first vowel: *i-uta* → *iyuta* ‘pound (grain)’, *u-osurpa* → *uwosurpa* ‘throw one another away, divorce’ (Chiri 1942: §15,1) — the same glide insertion seen in *imi-yehe* and *ruwe-he* above.

The *n*-final clitic *an=* assimilates to a following sonorant: *an=ye* ‘we say’ is pronounced *ayye*, *an=wante* ‘we know’ *awwante* (Chiri 1942: §16,3); before *s* the nasal dissimilates to *y* (*ay-sike* from *an=sike* ‘my load’), and before the desiderative the sequence *an=kor rusuy* surfaces as *an=kon rusuy* (§5.4). The corpus spelling *an=yee* ((Kitahara et al. 2003)) is morphophonemic; we take the phonetic value to be [ayyee], following Chiri.

5.4 r-alternations and the two kinds of -rV stems

Since *r* is banned from codas, a stem-final **r* surfaces with an echo of the preceding vowel: *koro* ‘have’, *nukara* ‘see’, *mukara* ‘axe’, *kisaru* ‘ear’ for Hokkaidō *kor*, *nukar*, *mukar*, *kisar*; less commonly the *r* joins the general coda neutralisation and yields *h*, as in *utah* beside *utara* ‘people’ and *mahnekuh* ‘woman’ ((Chiri 1942: §16,4); (Dal Corso 2024: 64)). Word-internally the old clusters survive as *h* + stop: *ohta* ‘at’ from *or ta*, *unahpe* ‘aunt’ (Tangiku 1998: §3-2).

Hattori took the echo vowel to be a real, historically added vowel, so that the Sakhalin verbs end in vowels and Hokkaidō preserves the older shape; Chiri treated it as an automatic repetition of the preceding vowel. Tangiku (1998) shows that for the West coast Chiri was right at the level of the underlying form: verbs in *-rV* fall into two classes that behave differently under suffixation. **Class I** verbs are those whose Hokkaidō cognates end in coda *r* — *koro* ‘have’ ~ *kor*, *nukara* ‘see’ ~ *nukar*, *mokoro* ~ *mokor*: they are consonant-final at the deeper level, and their surface vowel is the epenthetic echo, inserted only to open the syllable. **Class II** verbs are vowel-final in Hokkaidō as well — *tura* ‘accompany’, *poro* ‘be big’, *kira* ‘flee’ — and their final vowel is lexical, part of the stem itself. Nothing in the citation forms tells the two apart; only their behaviour under suffixation does.

Class I (underlying /-r/): -te, with <i>r</i> → <i>n</i>		Class II (vowel-final): -re	
<i>koro</i> ‘have’	<i>konte</i> ‘give’	<i>poro</i> ‘be big’	<i>porore</i> ‘enlarge’
<i>nukara</i> ‘see’	<i>nukante</i> ‘show’	<i>kira</i> ‘flee’	<i>kirare</i> ‘let escape’
<i>kara</i> ‘make’	<i>kante</i> ‘have made’	<i>tuuri</i> ‘stretch’	<i>tuurire</i> ‘stretch (tr.)’
<i>mokoro</i> ‘sleep’	<i>mokonte</i> ‘put to sleep’	<i>tura</i> ‘accompany’	<i>turare</i> —

Table 14: Causative -te/-re with -rV verbs (Tangiku 1998: §4)

The diagnostic environments are two. First, the causative: *-te* attaches to consonant-final stems and *-re* to vowel-final ones, and Class I *koro*, *nukara*, *kara*, *mokoro* select *-te* with the regular shift of *r* to *n* before it (*kor-te* → *konte*), while Class II *poro*, *kira*, *tuuri* select *-re* like any vowel-final stem (Tangiku 1998: §4). Lexicalised causatives preserve the same treatment: *cinupunte* ‘ennoble’ from *ci-nupuru-te* (Tangiku 1998: §4). Second, the desiderative *rusuy* and *rayki* ‘long to’: after the same four verbs the two *rs* dissimilate (*kon rusuy*, *nukan rusuy*, *mokon rayki*), whereas *tura rusuy*, *koore rusuy* keep their vowels (Tangiku 1998: §3-4).

- (7) *hemata e=kon rusuy hii*
 what 2SG.A=have DES Q

‘what do you want?’ — Murasaki 1979; Tangiku 1998: 226; West Sakhalin, Rayciska

- (8) *nean mahpoo-ho nukan rusuy kusu*
 that daughter-POSS see DES because

‘because he wanted to see that daughter of theirs...’ — Murasaki 1976; Tangiku 1998: 226; West Sakhalin, Rayciska

- (9) *tama in=konte*
 bead 1PL.EXCL.O=give

‘you gave me a bead’ — Dobrotvorskij 1875; Sakaguchi 2021; West Sakhalin, West coast, central (recorded 1867–1872)
 konte from kor-te already in the 1860s data; in= is the exclusive object marker otherwise known from the East coast (Dal Corso 2024: 66–69).

The shape of the alternation is itself a dialect diagnostic, perhaps the sharpest single one the verbal morphology offers. Hokkaidō causativises its *r*-final stems with a third allomorph *-e* — *kor-e* ‘give’, *nukar-e* ‘show’ — so the cognate verbs there never develop the nasal (Chiri 1942: §55). Sakhalin lacks *-e* entirely: a Class I stem takes *-te* with *r* → *n* on both coasts, from Dobrotvorskij’s 1860s sentences and Sentoku Tarōji’s East-coast letters ((Tangiku & Ogihara 2001: Letter 1)) to Asai Take’s tales a century later. A *-rV* verb that yields a causative in *-nte* or a desiderative in *-n rusuy* is thereby identifiable as Sakhalin at a glance, and the pair *konte* : *kore* draws the isogloss between the two branches in a single word.

The rules are ordered: suffixation of *-te* and *rusuy* applies to the consonant-final stem, then the echo vowel is inserted, and only then do the person-number suffixes attach (Tangiku 1998: §7). Hence the plural of *koro* is *koro-hci*, with the post-vocalic allomorph *-hci* rather than post-consonantal *-ahci* (§5.5), and the desiderative likewise takes *-ahci* outside it:

- (10) *rayciska onne nean reekoh tumikoro e-ariki rusuy-ahci manuu*
 Rayciska to that very war.make APPL-come.PL DES-PL REP

‘they say they very much wanted to come and make war on Rayciska’ — Murasaki 1979; Tangiku 1998: 231; West Sakhalin, Rayciska

The plural suffix attaches outside *rusuy*, which otherwise takes no person marking — evidence that *rusuy* is a suffix rather than an auxiliary (Tangiku 1998: §6).

The deletion also has a sharp outer boundary. The causative and the desiderative belong to the verbal complex and reach the underlying stem, but the homophonous onset of the converb *teh* ‘and, having ...-ed’ does not trigger the rule: *koro teh* is never reduced to the expected sandhi outcome **koh teh*, although *r + t* yields *ht* word-internally (§5.5) (Tangiku 1998: 225). In the ordering just given the echo vowel is in place before clause-level material attaches, and is then protected. The corpus bears the prediction out without exception: *koro teh*, *mokoro teh* and *yaycisekoro teh* recur throughout the West-coast texts, and the reduced form never occurs — the affix/converb boundary is exactly where the rule stops.

- (11) *neeteh taa rawpoka taa imii sohkara koro teh taa mokoro manu*
 and.then EMPH underneath EMPH garment bedding have and EMPH sleep REP

‘and then she spread her garment underneath as bedding and slept’ — Murasaki 2001: text 46; West Sakhalin, Odasu (Asai Take)

koro keeps its echo vowel before the converb *teh* (not **koh teh*), though the same stem gives *konte* and *kon rusuy*; *mokoro* likewise stands in its full surface shape.

In Asai Take’s Odasu tales the same morphology recurs (*nukan rusuyahci* ‘they wanted to see’, (Murasaki 2001: text 52)), so the system holds across the West-coast dialects. A solitary counterexample is *e-hecinte* ‘let play there’ from *hecire* ‘dance, play’, the only verb in $-V_1rV_2$ with unlike vowels that takes *-te*; Tangiku connects it with the East-coast and Hokkaidō variant *heciri* (Tangiku 1998: §5-2). The choice of echo vowel is occasionally lexical rather than copying — *kisaru* ‘ear’ matches the Hokkaidō possessed stem *kisar-u* (Tangiku 2022: §2.4), and the possessed form *kisaru-hu* appears in the Yamada Hayo corpus ((Kitahara et al. 2003)).

5.5 Plural allomorphy and connected speech

The verbal plural suffix has the shape *-ahci* after consonants and *-hci* after vowels, including inserted echo vowels: *cis-ahci* ‘they wept’, *itak-ahci* ‘they spoke’, *tura-hci*, *koro-hci*, *ariki-hci* (Tangiku 1998: §6) (see Chapter

8). Furukawa reports a Rayciska tendency for stems in *u* to take *-ahci* with vowel replacement (*konupuru* → *konupurahci* ‘they like it’), which the Odasu speaker does not share (*yaykonopuru-hci*) – dialect or idiolect variation that the small corpus cannot resolve ((Furukawa 1967); (Tangiku 1998: note)).

- (12) *husko ohta oyanrurukotan manuyke sine horokewpomaci-hi tura eyaykotankhany- manu*
 old.time in Oyanruru.village.SG and one young.man wife- with APPL.REFL.village.PL have-REP
 POSS PL PL

‘long ago there was the village of Oyanruru, and a man together with his wife had made it their home and lived there’ – Kitahara et al. 2003; Sakhalin (dialect not further specified), Yamada Hayo

Both allomorphs in one sentence: *-hci* after the echo vowel of *koro*, *-ahci* after consonant-final *okay*; *maci-hi* is the possessed form of *mah* ‘wife’ (§5.1).

On the East coast the suffix shows further erosion: after monosyllabic open-syllable verbs the *h* may drop (*nu-ci* ‘they heard’ beside expected *nu-hci*), a process never found in the West-coast sources (Dal Corso 2024: 65–66); Chiri had already noticed the conditioning, assigning *-si -ci* to long-vowel stems and *-hsi -hci* to short ones (Chiri 1942: §41). The parallel deletion in the possessive (*kotanu* for *kotan-[h]u*) was noted in §5.2. Sentoku Tarōji writes *itaku=hci* ‘they say’ with an East-coast support vowel where the West coast has *itak-ahci* ((Tangiku & Ogihara 2001: Letter 1)) – we read his *-u-* as the same cluster-resolving epenthesis that Dal Corso posits for *-VhV*.

Connected speech is governed by a small set of assimilations and dissimilations, most of them already catalogued by Chiri (1942: §§14–16): *r* assimilates to a following *n* (*cise-kon-nispa* ‘master of the house’ from *koro*), *r + t* gives *tt* in Hokkaidō but *ht* in Sakhalin (*kapah tuki ~ kapara tuki* ‘thin cup’), *n* becomes *m* before labials (*pom mahpoo* ‘little girl’), and *n m* dissimilate to *y* before *s* (*amay-su* ‘rice pot’ from *amam-su*; East-coast *tuy sata* from *tum sa ta* ‘in the middle’, (Dal Corso 2024: 65)). Vowel contraction trims frequent collocations (*setoko* from *si-etoko* ‘ahead of oneself’), and *h* flanked by a consonant and a vowel deletes (*sinotoripi* from *sinot-horipi* ‘play-dance’) (Chiri 1942: §§14–15).

Reduplication, finally, is lexical rather than inflectional in the recorded material. Body-part and place nouns show fossilised doubling of a final root (*etu-puy* ‘nose’ : *etu-puy-puy* ‘nostril’, *supuy-puy* ‘anus’, (Chiri 1942: §19,1)), verb-stem reduplication marks plurality or iteration of the event ((Dal Corso 2024: 55–56)), and sound-symbolic doublets such as *purupuru* carry the refrains of Asai Take’s sung tales ((Murasaki 2001: texts 34, 54)). We find no productive grammatical reduplication in the corpus: plurality is carried by *-hci/-ahci* and suppletive stems (Chapter 8).

Part III · Word classes and the noun phrase

Chapter 6 · Word classes

The parts of speech of Sakhalin Ainu and the criteria that distinguish them.

Sakhalin Ainu has two large open classes, nouns and verbs, and a set of closed classes: postpositions, demonstratives and other adnominals, numerals, adverbs, conjunctive particles, sentence-final and discourse particles, and interjections. There is no separate class of adjectives: property concepts are intransitive verbs (§6.4). The first systematic classification along these lines is Chiri (1942), whose Sakhalin-centred grammar distinguishes predicators (his 用詞, covering both action and property words, and split by a single morphological diagnostic into complete and incomplete verbs, §6.3), nominals (体詞), adverbs, adnominals, and seven classes of particles (Chiri 1942: §§21, 65, 85, 92, 117); Murasaki's grammar of the Rayciska dialect operates with essentially the same inventory (Murasaki 1976). This chapter states the criteria that separate the classes and illustrates each with corpus material; the morphology and syntax of each class are treated in the chapters that follow, above all Chapter 7 and Chapter 12.

6.1 Criteria

Word-class membership in Sakhalin Ainu is established morphosyntactically, not notionally. The two tests with the sharpest results are person marking and case marking. A verb is a word that heads a predicate and is indexed for the person of its arguments by the prefixes *ku-*, *e-*, *eci-*, *an-* and their kin (Chapter 13); a noun is a word that heads a noun phrase, serves as argument of a verb, and combines with case postpositions (Chiri 1942: §65). Person prefixes attach to nouns as well, but there they mark a possessor on the possessed (“belonging”) form in *-hV/-VhV* (§7.1), not a clause participant. The possessive suffix itself does not separate the two classes: Sakhalin Ainu attaches it to verbs too, as a nominalizing device (Chiri 1942: §77), (Bugaeva 2016: 104–107). What does separate them is distribution: only nouns take case postpositions and head argument phrases, and only verbs take the full argument-indexing paradigm, number marking such as the plural *-hci*, and the mood and aspect particles of Chiri's first particle class (Chiri 1942: §§41, 117).

Class	Open?	Criteria
noun	open	heads argument NPs; combines with case postpositions; possessed form in <i>-hV/-VhV</i> with possessor prefixes
verb	open	heads predicates; argument person prefixes/suffixes; strict lexical transitivity; verbal number (<i>-hci</i> , stem suppletion)
postposition	closed	follows an NP; assigns a case-like relation; uninflected
demonstrative/adnominal	closed	precedes the noun; no person or number morphology; cannot be predicate
numeral	closed	prenominal like adnominals; nominalized with classifiers (<i>-p/-pe</i> , <i>-n</i> , <i>pis</i>)
adverb	closed (plus derivation)	modifies predicates; no inflection; many derived in <i>-no</i>
conjunctive particle	closed	clause-final; links the clause to a following clause
final/discourse particle	closed	sentence-final or phrase-final; evidential, illocutionary or information-structural meaning
interjection	closed	extra-syntactic; utterance-initial or free-standing

Table 15: Word classes and their leading criteria

6.2 Nouns

Nouns show no gender, no case inflection, and no obligatory number inflection (Chiri 1942: §§66–68); the plural markers *utah* ‘people’ and *-ahcin* are optional and semantically restricted (Chapter 8; (Sakaguchi 2020a), (Sakaguchi 2020b)). The one inflectional opposition every speaker commands is that between the conceptual form and the possessed form, *cise* ~ *cise-he* ‘house’ ~ ‘his/her house’, treated in depth in Chapter 7. A noun phrase is head-final: adnominal demonstratives, numerals, possessor nouns and modifying verbs all precede the head. In (1) the demonstrative *tan* ‘this’ introduces the NP, the head *nay* ‘river’ serves as possessor of the following possessed noun, and the whole stands as subject of a nominal question predicate.

- (1) *tan nay ree-he temana an-pe*
 this river name-POSS what exist-NMLZ

‘What is the name of this river?’ — Murasaki 2001: text 38; West Sakhalin, Odasu (Asai Take)

Within the class, four subclasses need to be distinguished. (i) Common nouns, as in (1). (ii) Locative (relational) nouns such as *onnayke* ‘inside’, *kaske* ‘top’, *empoke* ‘underside’, *sanke* ‘vicinity’: they follow an NP directly, they alone combine with the simple locative postpositions *ta*, *wa*, *ene*, and when person-marked they take the object prefixes rather than the possessor set Chiri (1942: §79), (Sakaguchi 2024: 20); see §7.5 and Chapter 11. (iii) Formal (bound) nouns such as *pe* ‘thing’, *hi* ‘place, time’, *ruwe* ‘trace’, *kusu*, which never head a free NP and have lost person marking (Chiri 1942: §80); they supply the language’s nominalizers and much of its clause-linking machinery (Chapter 20, Chapter 23). (iv) Personal pronouns, *kuani* ‘I’, *eani* ‘you’, *anoka* ‘we, I’, *cioka* ‘we’: a small closed set, used only for emphasis or disambiguation, since the person affixes carry the functional load (Chiri 1942: §§81, 90); see Chapter 9. Proper names stand somewhat apart from common nouns in that they form no possessed form (Chiri 1942: §77).

6.3 Verbs

The verb is the morphological centre of the language. Three properties define the class. First, person: a verb is indexed for its subject and object by affixes, third person being zero (Chiri 1942: §§22–23, 37). Second, strict transitivity: every verb stem is lexically specified as intransitive or transitive, and valency is changed only by overt morphology — causative *-te/-re*, applicative *e-/ko/-o-*, antipassive *i-* and the like (Chapter 14). Pairs

such as *ahun* ‘enter’ versus *ahunke* ‘put in’ or *maka* ‘open (something)’ versus *makke* ‘come open’ show the opposition lexically (Chiri 1942: §§28, 35). Third, number: a set of stems distinguishes singular and plural by suppletion or stem alternation (*oman* ‘go (one)’ ~ *paye* ‘go (several)’, *an* ‘exist (one)’ ~ *okay* ‘exist (several)’), and Sakhalin Ainu adds the plural suffix *-hci* to verbs of either class (Chiri 1942: §§35, 41); Chapter 8. Verbs have no tense inflection; temporal and aspectual meaning is carried by particles and periphrasis (Chiri 1942: §47) (Chapter 15). In (2) the letter-writer indexes the agent with the indefinite-person prefix *an=* and pluralizes with *=hci*.

- (2) *ponno ponno cep ukoyki an=e=hci*
 little little fish catch INDEF.A=eat=PL

‘We catch just a few fish and eat them.’ — Tangiku & Ogihara 2001: Letter 1; East Sakhalin, Otasan (Sentoku Tarōji)

The second property carries more weight than the bare statement suggests, for Chiri made it the organising axis of his classification. Predicators divide into *complete* verbs (完用詞), which make a clause without any object — the intransitives, including every property verb of §6.4 — and *incomplete* verbs (不完用詞), which cannot (Chiri 1942: §§27–28). Membership is read off a single morphological fact, not off translation: the marker Chiri treats as the first person singular — the AN-set, glossed INDEF in this grammar (Chapter 13) — is suffixed to a complete verb but prefixed to an incomplete one, *ahun-an* ‘I went in’ against *an-ahunke* ‘I put it in’ from the pair above (Chiri 1942: §28). Example (2) shows the prefixal half of the diagnostic on the incomplete *e* ‘eat’; the suffixal half appears in Chiri’s own minimal demonstration, where the antipassive has stripped the object from the incomplete *ku* ‘drink’:

- (3) *i-ku-an*
 ANTIP-drink-INDEF.S

‘I drank (something).’ — Chiri 1942: §33; Sakhalin (dialect not further specified)

The derived stem is objectless, hence complete, and the person marker moves to the suffix slot; underived *ku* ‘drink (it)’ takes the prefix.

As the example shows, completeness is not frozen in the lexicon: it tracks every change of valency, and the person marking switches sides with it. Operations that add an object — the applicatives *e-*, *ko-*, *o-* and the causatives — turn a complete verb into an incomplete one; operations that fill or suppress the object slot — antipassive *i-*, reciprocal *u-*, reflexive *yay-* and *si-* — turn an incomplete verb into a complete one (Chiri 1942: §33). Each of these derivations is described in Chapter 14, where the effect of every operation on completeness is noted; the diagnostic itself, with Chiri’s homophonous-stem pairs, is taken up again in §12.2.

The copula *ne* and the existentials *an* and *okay* are verbs by every test above, though of minimal valency. The copula even has a place of honour in the dichotomy: Chiri singles it out as the one underived verb whose obligatory second term is a complement rather than an object, so that it inflects with the prefixes of the incomplete class (Chiri 1942: §32). These verbs are described with nominal predication in Chapter 18, and the negative existential *isam* ‘not exist’ with negation in Chapter 17.

- (4) *cis-ahci yayne tani taa reera yuhke ike poro ahto ran manu*
 cry-PL while now EMPH wind be.strong and be.big rain fall REP

‘As they wept, the wind grew strong and a great rain fell.’ — Murasaki 2001: text 8; West Sakhalin, Odasu (Asai Take)

The particle *taa*, ubiquitous in Asai Take’s narration, is glossed EMPH throughout this grammar, following Tangiku (2022). The clause is repeated with intensification in the original; we quote the first occurrence.

Example (4) already shows the central fact of the next section: the property words *yuhke* ‘be strong’ and *poro* ‘be big’ are verbs, the first used predicatively, the second pronominally.

6.4 Property concepts are verbs

There is no adjective class. Chiri (1942: §21) assembles the argument for Ainu generally, with Sakhalin material: property words take the same person marking as verbs (*ku-poro* ‘I am/grow big’, *ku-pirika* ‘I am well’); they take the same supportive *ki* ‘do’ (*poro ka ki* ‘is indeed big’ like *oman ka ki* ‘does indeed go’); they form imperatives (*tunasno pirika!* ‘get well quickly!’); and they are aspectually verb-like, *poro* meaning ‘be big’ or ‘become big’ according to context. The corpus bears this out. Predicative use with a possessed-form subject:

- (5) *etu-wehe pirika teh taa orowa cise ohta makan manu*
 beak-POSS be.good and EMPH then house at go.up REP

‘Its beak healed, and then it went back up to its house.’ — Murasaki 2001: text 13; West Sakhalin, Odasu (Asai Take)

Attributive (prenominal) use, where the property verb precedes its head noun exactly as in (4) above:

- (6) *tani pirika cise tah kara manu*
 now be.good house this make REP

‘Now a fine house had been made of it.’ (lit. ‘now (it) made this (into) a fine house’) — Murasaki 2001: text 46; West Sakhalin, Odasu (Asai Take)

The prenominal use is not a special adjectival construction: any verb can stand before a noun under the same conditions, and the modification structure preserves the argument relations of the corresponding clause, as Kirikae shows for Hokkaidō noun phrases (Kirikae 1984: 105–110). *Pirika cise* ‘good house’ is thus structurally a minimal relative clause, ‘a house that is good’ (Chapter 21). Manner modification is derived: the adverbializer *-no* turns a property verb into an adverb (Chiri 1942: §64).

- (7) *pirika-no reske*
 be.good-ADV raise

‘(She) raised (the child) well.’ — Murasaki 2001: text 37; West Sakhalin, Odasu (Asai Take)

6.5 Closed classes

6.5.1 Postpositions and case particles Grammatical relations of place, goal, source, instrument and company are marked by postpositions, Chiri’s fourth particle class, which attach only to nouns (Chiri 1942: §122): *ta* and *ohta* ‘at, in’, *onne* ‘to’, *ene* ‘towards’, *orowa* ‘from; and then’, *ani* ‘with, by means of’, *tura* ‘together with’, *kaari* ‘along, through’, *poka* ‘through, by way of’, *pahno* ‘as far as’. Common nouns require *ohta* (historically *oro-ta*) where locative nouns take bare *ta* — the main distributional test separating the two noun subclasses (§6.2). The full inventory and its syntax are given in Chapter 11.

- (8) *neeteh tani taa caru-hu ohta taa oro-oo*
 and.then now EMPH mouth-POSS at EMPH there-put

‘And then he put them into his mouth.’ — Murasaki 2001: text 30; West Sakhalin, Odasu (Asai Take)

6.5.2 Demonstratives and other adnominals Adnominals (Chiri’s 連体詞) modify a noun and do nothing else: they head no phrase, take no person marking, and never serve as predicates (Chiri 1942: §92). The Sakhalin demonstrative set is proximal *tan* ‘this’, distal *tara* ‘that yonder’ (also *tara-an*), against Hokkaidō *toan*; Chiri cites *tara seta* ‘that dog’ from Pilsudski (Chiri 1942: §97). In the corpus of Asai Take 浅井タケ the anaphoric *neya* ‘that aforementioned’ picks up an established referent, the pronominal *tah* ‘this one’ refers to things, and *taa* serves chiefly as a discourse particle of staging (glossed EMPH). Demonstratives and numerals may stack, in that order, but two demonstratives may not (Chiri 1942: §94). *sine* ‘one’ doubles as an indefinite article-like marker, ‘a certain’ (Chiri 1942: §115), as in (10) below. See further Chapter 9.

- (9) *neya cise taa onne taa neya tunkasapahka ahunu wa isam manu*
 that house EMPH to EMPH that skull enter and not.exist REP

‘Into that house the skull went and disappeared.’ — Murasaki 2001: text 34; West Sakhalin, Odasu (Asai Take)

6.5.3 Numerals Numerals pattern with adnominals: *sine* ‘one’, *tu* ‘two’ (*tuh*), *re* ‘three’ (*reh*) and so on stand before a counted noun and cannot stand alone; to count in isolation they must be nominalized by a classifier, *-p/-pe* for things, *-n* for people, *pis* for ‘pieces’ (Chiri 1942: §113). The system is vigesimal (*hoh* ‘twenty’, (Chiri 1942: §108)), and Sakhalin speakers also used a decimal series borrowed through the Amur (“Santan”) trade, *sine kunkutu* ‘ten’, *sine tanku* ‘hundred’ Chiri (1942: §112), (Murasaki 2009), (Sakaguchi 2022). Chapter 10 gives the full system.

- (10) *sine yeesu maci-hi tura okay-ahci*
 one Yeesu wife-POSS with exist.PL-PL

‘There once lived a Yeesu with his wife.’ — Murasaki 2001: text 52; West Sakhalin, Odasu (Asai Take)

Yeesu: a stock character of Sakhalin folktales.

- (11) *niinuntuh pis asin- ike taa simah taa kosontokocuputeh taa muhruempokene ahunkemanu*
 woodentball piece come.out EMPH woman EMPH silk.robe wrap and EMPH pil- un- to put.in REP
 NMLZ low der

‘When the two wooden balls came out, the woman wrapped them in a fine robe and put them under her pillow.’ — Murasaki 2001: text 52; West Sakhalin, Odasu (Asai Take)

muhru empoke ene ‘to under the pillow’: locative noun *empoke* with the allative *ene* (§6.2, §7.5).

6.5.4 Adverbs Adverbs modify a predicate or a whole clause and are themselves uninflected: they take no person, number, mood or case morphology, which is what separates them from verbs and nouns, and they precede the element they modify rather than following it, which separates them from the particles. Chiri sets the adverb (副詞) up as a class of its own, distinct from both the nominals and the adnominals, and notes that the Sakhalin adverb lexicon diverges considerably from Hokkaidō (Chiri 1942: §§64, 88). The class has three morphological sources. (i) A small closed set of *underived* adverbs, mostly temporal, aspectual and degree words: *tani* ‘now’, *suy* ‘again’, *naa* ‘still, more’, *ramma* ‘always’, *reekoh* ‘very’, *ponno* ‘a little’. (ii) A large open set of *manner* adverbs derived from property verbs by the adverbializer *-no* (§6.4): *pirika-no* ‘well’ from *pirika* ‘be good’, *monas-no* ‘quickly, at once’ from *monas* ‘be quick’ (Chiri 1942: §64). Because any property verb can feed it, this is where the productivity of the class lies; the underived core stays small. (iii) *Reduplicated* and expressive adverbs, in which a base is doubled for intensity or distributed reading — *ponno ponno* ‘just a little, bit by bit’ (example 2 above, and the ‘skip a little’ of Asai Take’s *ponno ponno...cookoko*) — continuous with the lexical reduplication described in Chapter 5 (Morphophonology).

By meaning the inventory falls into a handful of classes. *Temporal* adverbs locate the event in time: *tani* ‘now’, ubiquitous in the Asai corpus as a narrative ‘now, at this point’, *tanto* ‘today’, *nuuman* ‘yesterday’, *simma* ‘tomorrow’ (*simma e=oman* ‘go tomorrow’, Asai Take). Since the verb has no tense (§15.1), these words carry much of the temporal location that inflection does elsewhere. *Aspectual and frequency* adverbs quantify over occurrences: *suy* ‘again’, the commonest of all, which restarts an episode in narrative; *naa* ‘still, yet’, which in Sentoku Tarōji’s letters marks a state as continuing (*naa...e=an* ‘you are still there’); and *ramma* ‘always, as ever’, likewise from the Sentoku letters (*ramma ka pirika* ‘are well as ever’). The completive notion ‘finish doing’ is carried not by an adverb but by the verb *hemaka* ‘finish’ (Chapter 15), the iterative-habitual by the conjunctive *koh* (Chapter 15). *Degree* adverbs scale the predicate: intensifying *reekoh* ‘very’, which the Karafuto wordlists label explicitly as an adverb ‘very, exceedingly’; diminishing *ponno* ‘a little’; and the quantifier-like *okore* ‘all’ (*okore oyra* ‘forget all’), the Sakhalin counterpart of Hokkaidō *opitta*, whose own reflex *opista* the corpus uses only in the Sentoku letters. *Manner* adverbs are chiefly the *-no* class, joined by the demonstrative-based *nah* ‘thus, so’ ‘in that way’ (*nah kii* ‘do so’) and *ene* ‘like this’. The deictic and directional adverbs — *teeta* ‘here’, *ene* and their kin — pattern with the demonstrative system and are treated under it (Chapter 9).

- (12) *suy niina kusu makan manu*
 again gather.firewood PURP go.up REP

‘Again she went up into the hills to gather firewood, it is said.’ — Murasaki 2001: text 12; West Sakhalin, Odasu (Asai Take)

suy ‘again’ stands clause-initially, scoping over the whole event; the purpose phrase *niina kusu* precedes the matrix verb *makan*.

- (13) *reekoh pirika haw nee manu*
 very be.good sound COP REP

‘It was a very fine sound, it is said.’ — Murasaki 2001: text 51; West Sakhalin, Odasu (Asai Take)

The degree adverb *reekoh* modifies the property verb *pirika* directly, immediately to its left.

Syntactically, adverbs are not arguments and take no postposition; they attach to the predicate rather than to a noun. The default position is immediately before the constituent modified — degree adverbs hug the property verb, as *reekoh pirika* in the second example, and manner *-no* adverbs sit just before their verb, as *pirika-no reske* ‘raised (the child) well’ in (7) — while temporal and aspectual adverbs with clausal scope favour clause-initial position, as *suy* and *tani* repeatedly do in narrative. Two neighbouring devices are not part of this class. Negative polarity is not adverbial: it is built on the *ham*-family negators (*hannehka*, *hanki*, the proclitic *ham=*) and the negative existential *isam*, all treated in Chapter 17. And the scalar-additive *ka*, *kayki* and the restrictive *pateh* ‘only’ are clitic focus particles, placed after their host rather than before it (§24.2). The interaction of

adverb placement with the basic verb-final order of the clause is taken up in Chapter 19 (The simple clause and word order).

6.5.5 Conjunctive particles Clause linkage is managed by clause-final conjunctive particles, Chiri's third particle class (Chiri 1942: §120): sequential *teh* 'and' (the Sakhalin counterpart of Hokkaidō *wa*, (Chiri 1942: §121)) and *wa*, both glossed 'and' in this grammar; *ike* 'and, and then'; causal *kusu* 'because' and *renkayne* 'because of'; concessive *yahka* and *yahkayki* 'although'; conditional *anah* 'if', *ciki* 'if', *cikin* 'if'; habitual-correlative *koh* 'whenever'. *kusu* also marks purpose, and with a following existential builds an incipient progressive (*kayo kusu an* 'is calling', Chapter 15). The division of labour among these linkers is the subject of Chapter 23. Example (14) shows *anah* with the first person plural exclusive set; (15) shows *yahkayki* and *teh*.

- (14) *keh cioka ka ci-kii anah ci-ee*
 INTERJ 1PL.EXCL too 1PL.EXCL-do if 1PL.EXCL-eat

'Well then, if we do it too, we will eat.' — Murasaki 2001: text 41; West Sakhalin, Odasu (Asai Take)

The exclusive plural forms here refer to the speaker alone, a use of the plural regular in quoted speech in the Asai corpus (Sakaguchi 2024: 18).

6.5.6 Sentence-final and discourse particles The right edge of the Sakhalin sentence carries a rich set of particles. The reportative *manu* closes nearly every narrative sentence of the Asai corpus and marks the content as transmitted knowledge; it belongs to the evidential system described in Chapter 16 ((Dal Corso 2018); (Chiri 1942: §117) classes it among the mood particles). The quotative *nah* closes direct speech before a verb of saying or thinking (Chapter 22). The additive-scalar *ka* 'even, too' follows nouns and converbal clauses; the topic particles *anak* and *anakne*, prominent in Hokkaidō, are largely supplanted in the corpus by the construction with *neampe* (Chapter 24). Emotive and interactional particles include interrogative *hee*, softening *waa* (example 10 of Chapter 7), and *na* of assertion; Chiri's minimal pair is *seta he?* 'a dog?' against *seta un!* 'a dog!' (Chiri 1942: §128).

- (15) *nah yee yahkayki taa seta haw ne kara teh taa oman manu*
 QUOT say although EMPH dog voice COP make and EMPH go.SG REP

'Though she said so, he made his voice like a dog's and went off.' — Murasaki 2001: text 46; West Sakhalin, Odasu (Asai Take)

6.5.7 Interjections Interjections stand outside clause syntax: attention-calling *iineahsuy* 'say, hey', which opens requests and questions throughout the Asai corpus (so in example 10 of Chapter 7); *keh* 'well then' in (14) above; assenting *ee* 'yes'. They take no morphology and combine with no particle, which is what distinguishes them from adverbs.

Chapter 7 · Nouns and possession

Noun morphology, the possessive ("belonging") form, inalienable and alienable possession, and locative nouns. The Sakhalin Ainu noun inflects for neither gender nor case, and number marking is optional (Chiri (1942: §§66–68); Chapter 8). Its one pervasive inflectional category is possession: nearly every noun opposes a bare *conceptual form* to a *possessed form* in *-hV/-VhV*, *cise* 'house' ~ *cisehe* 'his/her house'. Kindaichi called the latter the "belonging form" (所属形); Chiri Mashiho 知里真志保, whose 1942 description remains the fullest, opposes the conceptual form (概念形) to person-marked forms built on a possessed stem (Chiri 1942: §§69–71). We write POSS for the suffix throughout. This chapter describes the formation and use of the possessed form, the syntax of possessor phrases, the inalienable–alienable divide, the locative noun subclass, and, in outline, derived nominals.

7.1 Conceptual and possessed form

Chiri introduces the opposition with *sik* ‘eye’ (Chiri’s morphophonemic citation; the West-coast isolation form is *sis* by coda neutralisation, Chapter 4): the conceptual form names the bare concept, while the possessed form anchors the referent to a possessor, third person being zero-marked Chiri (1942: §70).

form	gloss
<i>sik</i>	eye (conceptual form)
<i>siki(hi)</i>	his/her/their eye(s)
<i>e-siki</i>	your (sg) eyes
<i>eci-siki</i>	your (pl) eyes
<i>ku-siki</i>	my eyes
<i>ci-siki</i>	our (excl) eyes
<i>an-siki</i>	our (incl/indef) eyes

Table 16: Conceptual and person-marked forms of *sik* ‘eye’, after Chiri (1942: §70); his short *siki* corresponds to *siki-hi* in the corpus, cf. (1)

- (1) *nah kii yayne tani taa siki-hi naske manu*
 QUOT do while now EMPH eye-POSS open REP

‘As he kept doing so, his eyes opened.’ — Murasaki 2001: text 50; West Sakhalin, Odasu (Asai Take)

The discourse particle *taa* is glossed EMPH throughout; see Chapter 6.

7.1.1 Formation Vowel-final stems copy their final vowel after an inserted *h*: *sapa* ‘head’ ~ *sapa-ha*, *cise* ~ *cise-he*, *poo* ‘child’ ~ *poo-ho* (Chiri 1942: §72). Stems in long *ii/uu* take *-(y/w)ehe*, *imii* ‘garment’ ~ *imi-yehe* Chiri (1942: §72), (Bugaeva 2022: ch. 9, §4.1), and in the corpus *etu* ‘beak’ ~ *etu-wehe* (example 5 of Chapter 6); short *-i*, *-u* and the glide codas are the domain of a Sakhalin innovation, the generalized *-he*, taken up in the next subsection. Unlike Hokkaidō, which opposes short *tek-e* to long *tek-e-he*, Sakhalin uses only the long shape (Bugaeva 2022: ch. 9, §4.1).

Consonant-final stems take *-VhV*. The descriptive tradition leaves the choice of vowel to the lexicon, with only loose tendencies by final consonant Chiri (1942: §72), (Murasaki 1979), but Chiri’s later study of Ainu vowel harmony shows the choice to be largely rule-governed: tabulating some 160 noun stems, he found the suffix vowel agreeing in harmonic class with the last vowel of the stem, the classes being A (*a*, *u*), B (*o*) and C (*e*, *i*), and A and B vowels never combining within the word (Chiri 1952: 463–464). On this analysis *nan* ‘face’ ~ *nan-uhu*, *ram* ‘heart’ ~ *ram-uhu* and *kotan* ‘village’ ~ *kotan-uhu* select an A-class suffix vowel after A-class *a/u*, and *sik* ‘eye’ ~ *sik-ih* a C-class vowel after C-class *i* — predicted, not memorized. What is genuinely lexical is the residue where harmony licenses more than one vowel and a stem (or a sense) settles on one: *sar-a* ‘its tail’ beside *sar-i* ‘its reed-field’ (Chiri 1952: 463). We treat such pairs as a short list of inflection-class diacritics, comparable to thematic vowels in declension systems elsewhere: the harmony rule assigns the class, and only the harmony-internal choices need be listed per noun. Dal Corso takes the underlying suffix vowel to be *i* and the first V of *-VhV* to be an epenthetic repair, citing free variation of the type *kotanu* ~ *kotanh* in Pilsudski’s corpus (Dal Corso 2024: 65, fn. 9); on that view the harmonic agreement states which quality the repair copies in, and the two analyses compose rather than compete.

Stems whose final stop has weakened to coda *h* resurface it before the suffix: *teh* ‘hand’ ~ *tek-ih*, *otoh* ‘hair’ ~ *otop-ih*, and *mah* ‘wife, woman’ ~ *maci-hi*, where restored *t* palatalizes before *i* (*kuci* from *kut-i* ‘his belt’, (Chiri 1942: §16)). The segmental alternations involved are treated in Chapter 5 and not repeated here; (2) shows the alternation live in one clause.

- (2) *orowa tani taa teh ani sitayki-hi neampe tek-ih kotahma*
 then now EMPH hand INS hit-NMLZ TOP hand-POSS stick

‘Then, when she struck it with a hand, her hand stuck fast.’ — Murasaki 2001: text 43; West Sakhalin, Odasu (Asai Take)

Conceptual *teh* under the instrumental, possessed *tekihi* as subject; on V-*hi* *neampe* ‘when...’ see Chapter 20.

7.1.2 The generalized *-he* Hokkaidō dialects treat nouns in final short *-i* and *-u* as special classes — *ni-ye(he)* ‘his tree’, with periphrasis usual for many *u*-stems — and give glide-coda nouns *-e(he)*, *haw-e(he)*. Sakhalin

levels all of these under a single *-he*. In monosyllables the levelling is complete: *nii* ‘tree’ ~ *ni-he*, *kuu* ‘bow’ ~ *ku-he* Chiri (1952: 463), (Chiri 1942: §72), with the regular shortening of the lengthened monosyllable before the suffix (Chapter 5), and likewise for the glide codas, *ay* ‘arrow’ ~ *ay-he*, *haw* ‘voice’ ~ *haw-he* (Chiri 1942: §72). The corpus offers no counterexamples in these classes, and we state the rule as exceptionless there.

Polysyllabic *-i/-u* stems instead admit both shapes: Chiri records *kampi-hi* beside *kampi-he* ‘his letter’ (Chiri 1942: §72), and the conservative copying variant was still the form Murasaki’s West-coast consultants produced, as (3) shows.

- (3) *ku-itanki-hi* *isam*
1SG.A-teacup-POSS not.exist

‘My teacup is not here.’ — Murasaki 2025; West Sakhalin, Rayciska

The polysyllabic *i*-stem *itanki* keeps the copying *-hi* where generalized *-he* would also be licit.

The generalization has thus run to completion in just the short-stem classes — monosyllables and glide-finals — and remains variable in polysyllables: a lexically gradual, structurally bounded change rather than an across-the-board replacement. A prosodic signature travels with the morphology: where Hokkaidō shifts stress rightward onto the suffix, *ru-é* ‘its track’, Sakhalin keeps it on the stem, *rú-he* (Chiri 1942).

7.1.3 Choosing between the two forms The conceptual form is used when the noun names a kind rather than an anchored referent: in compounding and incorporation (*teh-kotoro* ‘palm’, lit. hand-surface, (Bugueva 2022: ch. 9, §4.3); (Chiri 1942: §76)), in vocatives (*kakapo* ‘elder sister!’, (Chiri 1942: §76)), and in classifying predications such as (4), where ‘wife’ denotes a role.

- (4) *pirika maxnekux max ne koroči*
pirika mahnekuh mah ne koro-ci
be.good woman wife COP have-PL

‘They took beautiful women as wives.’ — Chiri 1942: §76; Piłsudski 1912: 69; East Sakhalin

Chiri remarks that possessed *maci* here would wreck the sentence: ‘they took his wife’.

Once a referent is anchored in discourse, the possessed form tracks it. In the tale of the sticking tree-root, the axe is introduced in the conceptual form under the instrumental and immediately resumed in the possessed form, now ‘his axe’:

- (5) *mukara ani sitayki ike mukara-ha kotahma*
axe INS hit and axe-POSS stick

‘He struck it with an axe, and his axe stuck fast.’ — Murasaki 2001: text 35; West Sakhalin, Odasu (Asai Take)

Tangiku, comparing parallel passages across the Asai corpus, finds exactly this division: first mention of an instrument may be conceptual or possessed, but the second mention within the scene is reliably possessed (Bugueva 2022: ch. 9, §4.3). Where anchoring is irrelevant the conceptual form may stand even for owned items, so the choice is pragmatic, not automatic (Chiri 1942: §76).

7.2 The syntax of possession

Adnominal possession is head-marking and strictly head-final: the possessor NP stands before the possessed-form head, with no linking element — the type *aynu cise-he* ‘the man’s house’. Pronominal possessor phrases are usually omitted, since the prefix identifies them (Bugueva 2016: 104).

- (6) *sine acahcipo cise-he kaskepoka horokewpo oman-i ike wente manu*
one old.woman house-POSS onto.top.of man go.SG-NMLZ and destroy REP

‘The man climbed on top of an old woman’s house and wrecked it.’ — Murasaki 2001: text 46; West Sakhalin, Odasu (Asai Take)

- (7) *horokewpo imi-yehe anihi mii*
man garment-POSS herself wear

‘The man’s garment she put on herself.’ — Murasaki 2001: text 40; West Sakhalin, Odasu (Asai Take)

With no overt possessor and no prefix, the possessed form is third-person by default and is read from context, here ‘her husband’:

- (8) *nah taa yee teh taa hoku-hu koore*
 QUOT EMPH say and EMPH husband-POSS give

‘So saying, she gave it to her husband.’ — Murasaki 2001: text 41; West Sakhalin, Odasu (Asai Take)

7.2.1 Person-marked possession First- and second-person possessors are indexed on the possessed form by the same prefix set that marks transitive subjects: *ku-* 1SG, *ci-* 1PL exclusive, *e-* 2SG, *eci-* 2PL, *an-* indefinite Chiri (1942: §73), (Sakaguchi 2024: 19); Chapter 13. The corpus shows the set at work on nouns of every semantic type: *ku-cip-ih* ‘my boat’ (9), *ku-kema-ha* ‘my legs’ (10), *ku-nan-uh* ‘my face’ (Sakaguchi 2024: 19), *eci-ona-ha* ‘your father’ (11), *an-ona-ha* ‘my father’ (12).

- (9) *tan cíś ne ámpe kuáni ku-cíphi né*
tan cis neampe kuani ku-cip-ih nee
 this boat TOP 1SG 1SG.A-boat-POSS COP

‘This boat is my boat.’ — Pilsudski 1912: Nr. 16, lines 29–30; Sakaguchi 2024: 19; East Sakhalin

- (10) *iineahsuy ku-kema-ha ku-turi kusu e-aa ciki pirika-no aa waa*
 hey 1SG.A-leg-POSS 1SG.A-stretch because 2SG-sit if be.good-ADV sit FIN

‘Say, I am going to stretch my legs, so if you are going to sit, sit carefully.’ — Murasaki 2001: text 41; West Sakhalin, Odasu (Asai Take)

The same prefix *ku-* marks the possessor on *kema-ha* and the subject on *turi*.

- (11) *eci-ona-ha sikewente kusu kayo kusu an*
 2PL-father-POSS carry.heavy.load because call PURP exist.SG

‘Your father is calling because the load on his back is heavy.’ — Murasaki 2001: text 27; West Sakhalin, Odasu (Asai Take)

kayo kusu an, lit. ‘is at calling’, is the progressive periphrasis (Chapter 15).

In first-person narration the indefinite set *an-* serves as narrative first person, on nouns as on verbs (Chapter 13):

- (12) *an-ona-ha ekantasa ona-ha-hcin sukup-ahci ohta an oruspe nee manu*
 INDEF.A-father-POSS previous father-POSS-PL grow-PL at exist.SG story COP REP

‘It is a story from the time when my father’s fathers were growing up.’ — Murasaki 1976: 14; West Sakhalin, Rayciska (Fujiyama Haru)

Example (12) also shows the plural *-hcin* stacked outside the possessive suffix, a Sakhalin development without Hokkaidō counterpart: it pluralizes possessor or possessum according to context (Sakaguchi 2024: 110–115), (Sakaguchi 2020b); Chapter 8. In (13) it marks the several possessors of (severally possessed) heads.

- (13) *okore sapahka-ha-hcin menke-hci omantene taa cis-ahci kusu okay-ahci manu*
 all head-POSS-PL shave-PL and.then EMPH cry-PL PURP exist.PL-PL REP

‘They had all shaved their heads and were weeping.’ — Murasaki 2001: text 34; West Sakhalin, Odasu (Asai Take)

7.3 Inalienable possession

Body parts, kin terms and part-whole nouns are the core of the possessive system: in connected speech they occur overwhelmingly in the possessed form, and a bare conceptual *sapa* or *ona* is largely confined to compounds, vocatives and kind-readings of the type in (4) Chiri (1942: §76), (Bugaeva 2022: ch. 9, §4.2). The corpus matches this distribution: hands, legs, heads, eyes, beaks, names, voices, fathers, mothers, sisters, wives, husbands and children appear as *tek-ih*, *kema-ha*, *sapa-ha*, *siki-hi*, *etu-wehe*, *ree-he*, *maci-hi*, *hoku-hu*, *poo-ho* in example after example above. At the other pole, nouns for natural phenomena — *cuh* ‘sun’, *ah*to ‘rain’, *atuy* ‘sea’, *wahka* ‘water’ — normally stay conceptual, there being no possessor to anchor; Chiri notes that even these can be coerced into a possessed form when a god speaks of *an-kanto-ho* ‘our own sky’ (Chiri 1942: §76), and Murasaki’s consultants allowed *ku-atuy-ehe* ‘my sea, the sea where I live’ (Bugaeva 2022: introduction). Inalienability is thus a strong statistical norm rather than a formal class boundary.

7.4 Alienable possession and the koro strategy

Here Sakhalin Ainu departs sharply from Hokkaidō. In Hokkaidō the suffixal possessed form is restricted to inalienables, and alienable possession requires the relative-clause periphrasis with *koro* ‘have’: *a-kor mosir* ‘our land, lit. the land we have’. In Sakhalin the suffixal construction has been generalized to nouns of all classes — houses (6), garments (7), boats (9), axes (5), villages and lands (*an-mosir-ihi* ‘our land’, (Murasaki 1979: 5)) — so that both inalienable and alienable relations are head-marked (Bugaeva 2016: 105); for the place of this appositive possessor-plus-possessed pattern within Ainu and the wider Pacific see (Bugaeva 2021). Even animate property is possessed suffixally: the ogre’s hunting dog is its *seta-ha* (Murasaki 2001: text 35), an example Tangiku singles out as characteristically Sakhalin (Bugaeva 2022: ch. 9, §4.2).

The *koro* periphrasis nonetheless survives beside the suffixal pattern. Chiri describes it for possessors with personhood: *an-koro mah* ‘my wife’, *nispa koro mah* ‘the chief’s wife’ (Chiri 1942: §76). The Sentoku letters use it for social relations and places, *an=koro nispa* ‘my dear sir’, *e=koro kotan* ‘your village’ (Tangiku & Ogihara 2001: Letter 1), and in the Asai corpus it alternates with the suffixal form within a single sentence:

- (14) *poo koro ike taa tani ohkayo poo-ho naa mahtekuh poo-ho naa koro*
 child have and EMPH now male child-POSS also female child-POSS also have

‘They had children — they had both a boy and a girl.’ — Murasaki 2001: text 39; West Sakhalin, Odasu (Asai Take)
 Conceptual *poo* under *koro* for ‘to get children’; possessed *poo-ho* once the children are referential.

Where both strategies are available the difference is one of construal. Murasaki’s consultant contrasted *ku-kotan-uhi* ‘my village’, said while in it, with *ku-koro kotan*, said of one’s own village from elsewhere — and put *ku-koro kotan* in the mouth of a god (Murasaki 1979: 83–84), (Bugaeva 2022: ch. 9, §4.2). We read the suffixal form as marking a referent anchored to its possessor in the speech situation, and *koro* as predicating ownership or control; the corpus alternations in (14) and in the letters are consistent with this division.

7.5 Locative nouns

Locative (relational) nouns — *onnayke* ‘inside’, *kaske* ‘top’, *empoke* ‘underside’, *sanke* ‘vicinity’, *okaake* ‘place behind, aftermath’, *etoko* ‘place ahead’, *uturuke* ‘interval’, *soyke* ‘outside’ — form a noun subclass with their own syntax: they follow an NP directly (*cise onnayke* ‘the inside of the house’), and only they combine with the simple locative postpositions *ta*, *wa*, *ene* Chiri (1942: §79), (Bugaeva 2022: ch. 9, §5.1). Their case syntax belongs to Chapter 11; here we note only how they participate in the possessive opposition. Before a locative postposition the conceptual form appears, as in (15); as argument of an ordinary verb the possessed form is required, as in (16) and (17) (Bugaeva 2022: ch. 9, §5.4).

- (15) *cise onnayke poka hemata humi-hi an manu*
 house inside through what sound-POSS exist.SG REP

‘From inside the house came the sound of something.’ — Murasaki 2001: text 4; West Sakhalin, Odasu (Asai Take)

- (16) *cise onnayke-he ekuhteka-hci*
 house inside-POSS darken-PL

‘They made the inside of the house dark.’ — Murasaki 2001: text 4; West Sakhalin, Odasu (Asai Take)

- (17) *cise-he onnayke-he ka reekoh muni-hi ka isan teh an*
 house-POSS inside-POSS even very rubbish-POSS even not.exist and exist.SG

‘Inside his house there was not even a speck of dust.’ — Murasaki 2001: text 25; West Sakhalin, Odasu (Asai Take)
 Possessed *cise-he* as possessor of possessed *onnayke-he*: possession recurses.

When a locative noun takes a personal possessor it selects the object prefixes, not the possessor set used by common nouns: *en-ka* ‘above me’ (Chiri 1942: §79), *i-osmak-e* ‘behind me’ in the narrative of Yamabe Yasunosuke 山辺安之助 (Sakaguchi 2024: 20). This object-like marking aligns the landmark argument of a locative noun with the object of a transitive verb, and we take it as the synchronic trace of their verb-like relationality.

7.6 Derived nominals

Nouns are derived from verbs chiefly with the formal nouns *pe* ‘thing, one’ / *-p* (*poro-p* ‘a big one’), *hi* ‘place, time, fact’, *kur* ‘person’ and *usi* ‘place’ (Chiri 1942: §§60–63); their syntax is treated with nominalization in Chapter 20. The nominaliser *-p* is given in Chiri’s morphophonemic citation; in West-coast isolation it

undergoes the regular coda-obstruent neutralisation of Chapter 4, as with *sik* ~ *sis* earlier in this chapter. The formal noun *kur* likewise appears in its morphophonemic shape, but its final *r* follows the separate coda-*r* alternation (West-coast *r* → *h* or an echo vowel, varying word by word) rather than the obstruent rule (Chapter 4). Distinctively Sakhalin is the extension of the possessive suffix itself to verbs: a clause may be nominalized by *-hV* on its predicate, as in *sitayki-hi neampe* ‘when (she) struck it’ in (2), *e=ki-hi* ‘how you are faring’ in the Sentoku letters (Tangiku & Ogihara 2001: Letter 1), or main-clause *an-koro-ho* ‘I have (a grandchild)’ noted already by Chiri (Chiri 1942: §77). Bugaeva analyses these V-POSS forms as genuine nominalizations and as the innovation that gave Sakhalin Ainu its distinctive non-finite clause types (Bugaeva 2016: 105–110). That they are nominal in syntax and not merely in shape is shown by their argument privileges: a V-POSS clause can stand as the subject of an evaluative or veridicality predicate, as in (18), where the nominalized clause is what is judged untrue.

- (18) *ku-ye-he* *sunke*
 1SG.A-say-NMLZ be.untrue

‘What I said is untrue.’ (lit. ‘my saying is a lie’) — Murasaki 1979: 95; Bugaeva 2016: 108; West Sakhalin, Rayciska

The V-POSS construction and its uses return in Chapter 20 and Chapter 16.

Chapter 8 · Number

Nominal plurality with *utara* and *-ahcin*, associative plurals, and number in the verb.

Sakhalin Ainu nouns do not inflect for number. *seta* ‘dog’ is ‘a dog’ or ‘dogs’, *aynu* ‘person’ ‘a person’ or ‘people’; Kindaichi’s sketch already pairs *sine aynu* ‘one person’ with *wan aynu* ‘ten people’, the noun unchanged (Sakaguchi 2020b: 2), and Chiri states flatly that number is not a grammatical category of the noun (Chiri 1942: §67). Plurality is nonetheless richly expressible, by an interlocking set of optional devices: the plural word *utah* ‘people’ ~ *utara*, the suffix *=ahcin* on possessed nouns, number-distinguishing demonstratives (Chapter 9), suppletive and alternating verb stems, and the verbal plural clitic *-hci*. None of these pluralises a singular; each makes overt a plurality already understood, and all may be left off. This chapter describes the nominal markers first, then number in the verb, and ends with numeral phrases and the mismatches that the optionality of the whole system produces. The chief modern studies are Sakaguchi (2020a) on *utara*, Sakaguchi (2020b) on *=ahcin*, and chapters 5–8 of Sakaguchi (2024) on verbal number; we follow them closely and add corpus observations from the West-coast tales of Asai Take (浅井タケ).

8.1 The plural word *utah* ~ *utara*

West-coast *utah* ‘people’ ~ *uta*, East-coast *utara*, corresponds regularly to Hokkaidō *utar* (coda *r* → *h* ~ *rV*, Chapter 4). Three uses must be kept apart (Sakaguchi 2020a: 188). First, it is an ordinary noun ‘kin, companions, people’, with the possessive form *utari(hi)* (Chapter 7). Second, placed after an adnominal, a noun or a relative clause it builds a plural noun phrase — *nea utara* ‘those people’, *aynu utara* ‘the people’, *kotan ohta okay utara* ‘the people who are in the village’ (Sakaguchi 2020a: 188) — and in this use it has become the language’s general plural marker. Third, unlike Hokkaidō *utar*, Sakhalin *utara* also occurs bare, with no modifier at all, as a pronoun-like plural ‘they, people, we, you’; this use was first described by Satō in 1987 and has been reanalysed by Sakaguchi (2020a), whom we follow.

8.1.1 *utara* as plural marker As a postposed plural marker *-utara* attaches to any noun denoting something animate (or treated as animate: Pilsudski’s texts have *ikoro utara* ‘treasures’, *kosonto utara* ‘silk robes’, (Chiri 1942: §67)), whether in the bare (conceptual) form or in the possessed form, and even to nouns that are already plural — Chiri cites *eciokay utara* ‘you-PL-people’ with the 2PL pronoun (Chiri 1942: §67), and Yamabe Yasunosuke (山辺安之助) writes *anokay-utara* ‘we’ beside the commoner *anokay-ahsin* (Sakaguchi 2020b: 6). With demonstratives the marker interacts with the number distinction of the demonstrative itself: *nea seta* ‘that dog’, *nea seta-utar(a)* ‘those dogs’, *neeroh nuso-utara* ‘those sledges’ (Chiri 1942: §67). In Asai Take’s tales the combination of anaphoric demonstrative, noun and *utah* is a narrative staple:

- (1) *neya kamuy utah renkayne wahka ohta hecire-hci wa okay-ahci manu*
 that god PL plentifully water in dance-PL and exist.PL-PL REP

‘Those gods, a great crowd of them, were dancing in the water.’ — Murasaki 2001: text 46; West Sakhalin, Odasu (Asai Take)

Semantically *-utara* covers both additive plurality (‘more than one X’) and associative plurality (‘X and those with X’) (Sakaguchi 2020b: 13–14). The associative reading is forced when the host is a personal name:

- (2) *Eekichi-utara hopihsanne makap-ahsi*
 Eekichi-PL beach.from go.uphill.PL-PL

‘Eekichi and his party came up from the beach.’ — Sakaguchi 2020b: 14; East Sakhalin

From Yamabe Yasunosuke's autobiography (Yamabe & Kindaichi 1913: 75). The plural stem *makap* and the clitic *-ahsi* (§8.3) agree with the associative group.

- (3) *Muramatsu Yoshino-utara naa eramusinne-hsi*
 Muramatsu Yoshino-PL too be.relieved-PL

‘Muramatsu and Yoshino, too, were relieved.’ — Sakaguchi 2020b: 17; East Sakhalin

Yamabe & Kindaichi (1913: 183). The context shows that exactly the two named men are meant: an exemplative use of the associative plural (‘the ones named, as a set’), in Yamakoshi's terms.

In the last example no third party is included: *-utara* lists its hosts exhaustively, what Sakaguchi (2020b: 17) identifies (after Yamakoshi) as an exemplative associative. The marker also attaches to possessed forms — *an-cinkew-he utara* ‘my ancestors’ (Chiri 1942: §67) — so the often-repeated rule that *-utara* takes the bare form and *=ahcin* the possessed form ((Murasaki 1979: 85, 92)) describes a strong tendency, not a constraint; the real division of labour is semantic (§8.2).

8.1.2 Bare *utara* and its grammaticalization Bare *utara* is frequent: 109 tokens in Pilsudski's Materials alone, against four of the third-person pronoun *anihi* (Sakaguchi 2020a: 189). Its uses are three (Sakaguchi 2020a: 192–197). Generic *utara* refers to people at large, with no antecedent:

- (4) *keraykusu asin tan mosiri kes ta koramusine okay utara ki manu*
 thanks.to.that now this world end LOC be.at.ease exist.PL people do REP

‘Thanks to that, people now live at ease in this latter age of the world.’ — Pilsudski 1912: text 10, ll. 102–103; Sakaguchi 2020a: 188; East Sakhalin

okay ‘exist.PL’ is here nominalized, ‘living, life’, the object of *ki* ‘do’.

Anaphoric *utara* picks up a plural human antecedent and is in effect a third person plural pronoun:

- (5) *ay-yup-utari- sap haworoke-sin an manu apa tuyka ta utara ahuh manu*
 INDEF.A- go.down.PL voice.NMLZ-PL exist.SG REP door top LOC people enter.PL REP
 elder.brother-PL-POSS

‘There was the sound of my brothers coming down; they came in at the door.’ — Pilsudski 1912: text 21, ll. 87–88; Sakaguchi 2020a: 189; East Sakhalin

utara resumes *ay-yup-utari-hi* ‘my brothers’. On *hawor-oke-sin* see §8.2.

Third, within quoted speech *utara* is used deictically of people present in the speech situation: ‘we’ when the speaker is included, ‘you (pl.)’ when not (Sakaguchi 2020a: 193–197):

- (6) *nahteh orowa enciw onne ikoiranu ham utara ki kusu*
 then from.now human to fall.in.love NEG people do PURP

‘So from now on let us no longer fall in love with humans.’ — Pilsudski 1912: text 24, l. 219; Sakaguchi 2020a: 190, 194; East Sakhalin

Fox brothers deliberating among themselves: *utara* includes speaker and hearers.

- (7) *utara urayki rusuy kusuneyke urayki yahka pirika*
 people fight DES if fight although be.good

‘If you want to fight it out, you may fight.’ — Pilsudski 1912: text 12, ll. 155–156; Sakaguchi 2020a: 196; East Sakhalin
 An old man to his visitor, of the visitor and the old man's son: *utara* = the hearer's party, excluding the speaker.

Sakaguchi takes the second-person use to be honorific avoidance — the addressee is not named directly, exactly as a wife addresses her husband with third-person *anihi* (Chapter 9) — since the plain 2PL pronoun *ec(i)okay* was available and is attested in brusque commands (Sakaguchi 2020a: 196–197). The deictic uses are not confined to the east coast; Murasaki's west-coast conversation recordings have bare *uta* of one's own circle:

- (8) *sake naa uta i-kuu-re*
 wine too people INDEF.O-drink-CAUS

‘They (everybody) gave me wine to drink, too.’ — Murasaki 1976: 77; Sakaguchi 2020a: 193; West Sakhalin, Maoka
i- is the indefinite object prefix in its first-person use (Chapter 13).

How far has bare *utara* grammaticalized? Sakaguchi's syntactic evidence points to a clitic-like element rather than a free pronoun. In negative clauses it stands between the negative adverb and the verb — *ham utara ki* ‘they/we do not do’ — where independent pronouns may not stand (*kuani han ku-nu* ‘I did not hear’, never **han kuani ku-nu*); it strongly prefers the position immediately before the verb, after an overt object; and it is never itself topicalised with *neanpe* (Sakaguchi 2020a: 190–191):

- (9) *an-matak-ih* *nah ye yahka ham utara nu-hci*
 INDEF.A-younger.sister-POSS QUOT say although NEG people listen-PL

‘Though my younger sister said so, they would not listen.’ — Pilsudski 1912: text 12, ll. 58–59; Sakaguchi 2020a: 190; East Sakhalin

Yet an object may exceptionally intervene between *utara* and the verb (*utara kam eimeh* ‘she dealt them out meat’, (Sakaguchi 2020a: 191)), so the attachment is not yet morphological. Diachronically the anaphoric use is best derived by ellipsis of the modifier from the *X-utara* construction, and the generic and deictic uses from the noun's own sense ‘companions, kin’ (Sakaguchi 2020a: 198–199). We accordingly describe bare *utara* as a defective plural pro-form on its way from noun to agreement-like clitic — functionally parallel to the AN-series of Hokkaidō Saru, but a Sakhalin development out of native material. It refers only to animates, and the verb beside it may, but need not, carry the plural clitic *-hci*: Satō found *-hci* on about half of the predicates of bare *utara* in the east-coast folktales (Sakaguchi 2024: 97), and in our count of Asai Take's tales a *-hci*-marked verb appears in roughly six out of seven sentences containing a noun phrase with *utah* — usual, never obligatory.

8.2 The nominal plural =ahcin

The second nominal plural is the suffixal clitic =ahcin, with allomorphs =hcin after vowels (the usual realisation, since its hosts end in the possessive suffix *-hV*), =ahcin after consonants, and southern and west-coast variants in *s* (=a)hsin, =sin, =cin (Sakaguchi 2020b: 4–5). Nothing like it exists in Hokkaidō; within Sakhalin it has been recognised since Pilsudski, who called *cin* ~ *sin* a plural sign typically of inanimates (Pilsudski 1912: 36), while Kindaichi mistook it for an emphatic particle — refuted by Chiri, who showed that it stands before case particles, where no such particle can stand (Chiri 1942: §67). Its distribution is rigidly restricted: it attaches only to the possessive (“belonging”) form of nouns, never to the bare conceptual form and never to proper names (Sakaguchi 2020b: 5–8, 13). Hattori's Rayciska paradigm shows the system (Hattori 1961: 6–7):

form	gloss
<i>seta</i>	‘dog(s)’ — number-neutral
<i>seta-ha</i>	‘his/her/their dog (one)’
<i>seta-ha=hcin</i>	‘his/her/their dogs’
<i>ku-seta-ha</i>	‘my dog (one)’
<i>an-seta-ha</i>	‘our dog — one shared, or one each’
<i>an-seta-ha=hcin</i>	‘our dogs’

Table 17: Number in the possessed noun (Rayciska; Hattori 1961: 6–7)

Within the possessed paradigm, then, number marking is close to obligatory: a possessed noun without =ahcin is singular — or distributively singular, one per possessor, as *an-seta-ha* shows. The same distributive singular

is textually attested: Yamabe writes *an-tek-i naa uk-ahsi* ‘they took our hands (each one hand)’ (Sakaguchi 2020b: 6), and Hattori elicited:

- (10) *anokay-ahcin an-tek-ih i ohta ama yan*
 INDEF-PL INDEF.A-hand-POSS in put IMP.PL

‘Put it into our hands (into each one’s hand).’ — Hattori 1961: 7; Sakaguchi 2020b: 5; West Sakhalin, Rayciska

Elicited by Hattori from his Rayciska consultants; *tek-ih i* remains unmarked because each person’s single hand is meant.

The suffix is equally at home in the west-coast narrative corpus, on kin terms, body parts and belongings:

- (11) *kotan-ta sine acah-cipo taa mici-hi-hcin tura an manu*
 village-LOC one old.woman there grandchild-POSS-PL with exist.SG REP

‘In a village an old woman lived together with her grandchildren.’ — Murasaki 2001: text 51; West Sakhalin, Odasu (Asai Take)

The next sentence specifies re horokewpo ‘three young men’; an agrees with the singular subject.

Semantically Sakaguchi (2020b: 8–9, 16–17) shows that *=ahcin* expresses additive plurality only — sets of like objects: siblings, fingers, shoes, pots. This is why it never attaches to personal names, whose referent is unique, and it is the key to the contrast with *-utara* drawn by Murasaki from her Fujiyama Haru materials: *ku=mici-hi-hcin* ‘my grandchildren’ (each one my grandchild) against *mit-utara* ‘the grandchildren (and such small fry)’, where the associative reading is available (Sakaguchi 2020b: 15–16). A marginal associative residue does exist: *an-(h)ok-ho=cin*, literally ‘my husbands’, refers in its tale to the husband and his younger brother, and Chiri glossed precisely this type ‘the husband and his companion’ (Sakaguchi 2020b: 8–9). In Asai Take’s tales we likewise find *aaca-ha-hcin* ‘her father and his household’ rendered associatively by Murasaki; we analyse such kin-term tokens as the same marginal associative extension, expected on relational nouns whose referents come in households, and keep the additive characterisation for the core. One further host type is the evidential nominalizers *hawhe(he)* ‘voice’, *humi(hi)* ‘sound, feel’, *siri(ihi)* ‘appearance’, *ruu(he)* ‘trace’ (Chapter 16, Chapter 20): formally possessive forms, they take *=ahcin* when the nominalized event is plural — and then the verb inside the clause carries no *-hci*, the two markers being in complementary service (Sakaguchi 2020b: 9–12):

- (12) *nispa-utara pinaponne u-moymoye hawhe-sin an manu*
 chief-PL secretly RECP-shake.awake voice.NMLZ-PL exist.SG REP

‘There came the sound of the chiefs secretly shaking one another awake.’ — Pilsudski 1912: text 21, l. 57; Sakaguchi 2020b: 10; East Sakhalin

- (13) *nea nea oyas-utah tani ruwes an okaakara makah siri-hcin an manu*
 that that demon-PL now shore.path along go.uphill.PL appearance.NMLZ-PL exist.SG REP

‘And then those demons were just coming up along the path from the shore.’ — Kitahara et al. 2003: 84; Sakaguchi 2020b: 12; Sakhalin (dialect not further specified), Yamada Hayo

Outside these types only the reciprocal verb *uriwahne* ‘be siblings to one another’ takes the suffix, in nominal use: *uriwahne-hcin* ‘brothers and sisters’ (Hattori 1964: 42), (Sakaguchi 2020b: 12–13). The element is formally similar to the verbal plural *-hci* of §8.3, and Chiri treated the two as one morpheme (Chiri 1942: §41); with Sakaguchi (2024: 103) we keep them apart synchronically — their hosts and shapes differ — while accepting the historical connection.

8.3 Number in the verb: the plural clitic *-hci*

Verbs mark number twice over: a small set of stems alternates or supplements for number (§8.4), and any finite verb may take the plural clitic *-hci* (*-ahci* after consonants; east coast also *-(a)hsi*, *-si*, west coast south *-cin*), whose form rules Chiri established (Chiri 1942: §41) and whose syntax Sakaguchi (2024: 77–102) has now clarified. The clitic attaches outside the person suffixes (*paye-an=ahci* ‘we went’), follows even auxiliaries, and disappears under nominalization, exactly like person marking — Sakaguchi therefore ranks it with the clitic-like person markers rather than with derivational plural stems (Sakaguchi 2024: 79–80, 102):

- (14) *usaanpe emuyke an-ruura hemaka-hsi*
 various.thing all INDEF.A-carry finish-PL

‘We finished carrying all the goods.’ — Sakaguchi 2024: 80; East Sakhalin

Yamabe & Kindaichi (1913: 180); *-hsi* cliticises to the auxiliary *hemaka*, outside the lexical verb.

What does it index? Murasaki listed it as an optional third person plural (Murasaki 1979: 49–51), but it co-occurs with overt person markers of any person, so it is a number marker, not person agreement — the conclusion Nakagawa too reaches, taking *-hci* to be an optionally deployed marker of plurality rather than part of the person paradigm (Nakagawa 2024: 151–153) (Chapter 13). Its controller is a core argument: normally the subject, but Pilsudski already glossed *an-kuyra-si* ‘I stole up on them’ as plural “because the object (the gods) is in the pl.” (Pilsudski 1912: 40–41). Sakaguchi’s synthesis of Hattori’s elicitation and the texts is that *-hci* supplements person marking wherever person affixes leave plurality of a human subject or primary object unexpressed: it is usual with plural third persons (zero-marked) and with the formally ambiguous indefinite series, and it is not used where *eci-* 2PL or the CI-series already settle the number (Sakaguchi 2024: 97–102). The controller is animate; plural inanimate subjects take a plural stem but not *-hci* (see example (20) below). In narrative it is pervasive:

- (15) *ehah-taa monimahpo utah hecire-hci wa sap-ahci kusu ikii*
 lily.root-dig girl PL dance-PL and come.down.PL-PL PURP do

‘The girls who went digging lily roots will come dancing down.’ — Murasaki 2001: text 5; West Sakhalin, Odasu (Asai Take)

Relative clause *ehah-taa* + *utah*; both verbs carry *-hci*, and *sap* is the plural stem (§8.4).

8.3.1 Plural imperatives Commands to more than one addressee add *(y)an* to the verb; after a consonant the *y* drops Murasaki (1979: 30), (Sakaguchi 2024: 48). The particle is the same element that marks second-person plurality in the 1↔2 interaction forms (Chapter 13):

- (16) *iineahsuy monimahpo uta tan pon horokewpo tura-yan kanne sap-an*
 hey girl PL this young man accompany-IMP.PL while go.down.PL-IMP.PL

‘Now, girls, go down together with this young man.’ — Murasaki 2001: text 5; West Sakhalin, Odasu (Asai Take)

The old woman addresses several girls; on *sap* the imperative surfaces as *-an*.

- (17) *eci-sapahka kotahma-re-yan eci-maci-hi-hcin ahun kusu ikii*
 2PL.A-head attach-CAUS-IMP.PL 2PL.A-wife-POSS-PL enter.SG PURP do

‘Stick your heads back on! Your wives are about to come in.’ — Murasaki 2001: text 5; West Sakhalin, Odasu (Asai Take)

eci-sapahka stays unmarked — one head per addressee, the distributive singular of §8.2 — while *eci-maci-hi-hcin* takes *=hcin*; on singular *ahun* with a plural subject see §8.5.

8.4 Suppletive and alternating verb stems

A closed set of intransitive verbs of motion, posture and existence distinguishes number in the stem, by suppletion or by alternation of the final consonant (Chiri 1942: §34–35). Sakhalin keeps the inherited core but has fewer pairs than Hokkaidō: the dialect dictionary records 21 singular–plural pairs for Rayciska against 42 for Saru, and the productive Hokkaidō plural suffix *-pa* survives only in a handful of transitive pairs (*tuye/tuypa* ‘cut’ type, (Murasaki 1979: 42)) (Sakaguchi 2024: 122–124). The table lists the pairs we can verify in Sakhalin sources; every plural stem in the right-hand column is also attested in the Asai Take corpus.

meaning	singular	plural	attestation
‘exist, be at’	<i>an</i>	<i>okay</i>	Chiri (1942: §35); Murasaki (2025: 17)
‘go’	<i>oman</i>	<i>paye</i>	Chiri (1942: §35); Murasaki (2025: 21–22)
‘come’	<i>eh</i>	<i>ariki</i>	Chiri (1942: §35); Murasaki (2025: 21)
‘sit’	<i>a</i>	<i>rok</i>	Chiri (1942: §35); <i>rokahci</i> , Murasaki (2001)
‘get up’	<i>numa</i>	<i>payki ~ payke</i>	Sakaguchi (2024: 122)
‘stand’	<i>as ~ hotari</i>	<i>roski</i>	Chiri (1942: §31); Sakaguchi (2024: 122)
‘take’ (tr.)	<i>uk</i>	<i>uyna</i>	Sakaguchi (2024: 122); <i>uynahci</i> , Murasaki (2001)
‘go out’	<i>asin</i>	<i>asip</i>	Chiri (1942: §35); Sakaguchi (2024: 123)
‘enter’	<i>ahun</i>	<i>ahup</i>	Chiri (1942: §35); Sakaguchi (2024: 123)
‘go down (seaward)’	<i>san</i>	<i>sap</i>	Chiri (1942: §35); Sakaguchi (2024: 123)
‘go uphill (inland)’	<i>makan</i>	<i>makap</i>	Chiri (1942: §35); Sakaguchi (2024: 123)
‘ascend’	<i>rikin</i>	<i>rikip</i>	Chiri (1942: §35); <i>rikipahci</i> , Murasaki (2001)
‘go up from the sea’	<i>yan</i>	<i>yap</i>	Chiri (1942: §35); <i>yapahci</i> , Murasaki (2001)

Table 18: Number-paired verb stems attested in Sakhalin Ainu

The *-n/-p* rows are a regular coda alternation confined to verbs of oriented motion; the plural coda surfaces as *h* before a consonant (*makah sirihcin* above; Chapter 4). The corresponding transitives split the same way, with singular *-ke* against plural *-te* (*ahun-ke* ‘let one in’, *ahup-te* ‘let many in’) and *oman-te* against *paye-re* (Chiri 1942: §36) (Chapter 14). The Hokkaidō pair *rayke/ronnu* ‘kill’ is not a number pair in Sakhalin: Hattori’s dictionary marks the two as stylistic variants, and in the texts plural killing is expressed by *rayki-hci* (Sakaguchi 2024: 123) — in Asai Take’s tales *raykihci* occurs 23 times, *ronnu* never. Asai’s plural of ‘go’ is consistently *paye(hci)*; a form *payeh* with final *h* is unattested in the corpus.

The choice of stem follows grammatical rather than referential number. The CI- and AN-series select plural stems even for one referent (Chapter 13); conversely the plural stem with *-an* is the normal ‘we’:

- (18) *keh anoka tu-aynu-ne sap-an*
 well INDEF two-person-COP go.down.PL-INDEF.S

‘Well then, let the two of us go down.’ — Murasaki 2001: text 30; West Sakhalin, Odasu (Asai Take)

8.5 Numeral phrases, optionality, and mismatches

Numerals modify the unmarked noun directly — *sine aynu* ‘one person’, *tu monimahpo* ‘two girls’, *re horokewpo* ‘three young men’ — or stand free with the classifier suffixes *-n* (humans) and *-p ~ -pe(h)* (things) (Chiri 1942: §105, §113), (Sakaguchi 2022) (Chapter 10). A numerally quantified noun normally takes no plural marker, and since Kindaichi it has been reported that it also prefers the *singular* verb stem, the numeral making plurality explicit enough (Sakaguchi 2024: 131). The corpus bears this out:

- (19) *re monimahpo an ruwehe an manu*
 three girl exist.SG trace.NMLZ exist.SG REP

‘There were signs that three girls lived there.’ — Murasaki 2001: text 5; West Sakhalin, Odasu (Asai Take)

But it is a preference, not a rule. Sakaguchi (2024: 133–140) shows numeral and *poronno* ‘many’ phrases with either stem, and argues that the real conditioning is construal: a quantity taken as one bounded mass selects the singular stem, individuated members select the plural. A single sentence of Yamabe’s makes the contrast:

- (20) *aynu ruu poronno an wahkayki husko ruu pateh okay*
 person track much exist.SG although old track only exist.PL

‘There were many human tracks, but only old ones.’ — Sakaguchi 2024: 140; East Sakhalin

Yamabe & Kindaichi (1913: 108). The mass of tracks takes *an*; the old tracks, individually inspected and newsworthy, take *okay* — and no *-hci*, the subject being inanimate.

The same looseness runs through the whole system. In Asai Take's tales we find *tu monimahpo utah an manu* 'there were the two girls', a plural-marked noun phrase with the singular existential (text 51), and *ahun* 'enter (sg.)' predicated of the plural wives in example (17) above; conversely fully redundant strings like example (1), where demonstrative, *utah*, plural stem and *-hci* all co-occur. We analyse the situation thus: number in Sakhalin Ainu is everywhere optional and semantically driven — markers are recruited to make plurality (or its associative and distributive shadings) explicit where it matters to the discourse, and left off where it is recoverable, the only near-grammaticalised corners being the possessed-noun paradigm of §8.2 and the grammatical concord of person series with stem number. The genuinely open question is the division of labour between plural stem and *-hci* with third persons; our working hypothesis, after Sakaguchi (2024: 140), is that the stem tracks construal (mass vs. individuated) while *-hci* tracks animate participants, and texts in which inanimate plurals take *-hci* would falsify it.

Chapter 9 · Pronouns and demonstratives

Personal pronouns, the indefinite person, demonstratives, and interrogative–indefinite words.

Sakhalin Ainu is a pro-drop language: person is carried by the verbal affixes of Chapter 13, and the words discussed in this chapter — personal pronouns, demonstratives, interrogatives — are a small, partly defective class added for emphasis, contrast and deixis. Chiri devoted a famously deflationary section to them: pronouns take no case forms, no modifiers, head no compounds, and *ani oman* 'HE went' means no more than *oman* 'he went' — "the pronoun is mere reinforcement" (Chiri 1942: §90). The data below come from Chiri's west-coast survey (Chiri 1942: §90–§104), Hattori's Rayciska elicitation (Hattori 1961), Murasaki's Rayciska grammar and conversation course (Murasaki (1979); Murasaki (2025)), Sakaguchi's east-coast philology (Sakaguchi (2020a); Sakaguchi (2024: 26–56)), the Sentoku letters (Tangiku & Ogihara 2001), and our own counts over the corpus of Asai Take's (浅井タケ) tales.

9.1 Personal pronouns

Table 1 gives the forms attested in Sakhalin sources. Each pronoun belongs to one of the person series of Chapter 13 and obeys the personal agreement rule stated there: a pronoun co-occurs only with verbal affixes of its own series.

person	form	sources
1SG (KU-series)	<i>kuani</i>	Chiri (1942: §90); Murasaki (2001)
1 (CI-series; sg. reference)	<i>ciokay ~ cookay ~ cokay</i>	Chiri (1942: §90); Sakaguchi (2024: 29); Tangiku & Ogihara (2001)
1SG/1PL (AN-series)	<i>anoka(y) ~ anokay</i>	Chiri (1942: §90); Tangiku & Ogihara (2001); Murasaki (2001)
1PL (dedicated)	<i>anokay-ahcin</i> ; also <i>anokay-utara</i>	Hattori (1961: 6); Sakaguchi (2020b: 6)
2SG	<i>eani</i>	Chiri (1942: §90); Tangiku & Ogihara (2001); Murasaki (2001)
2PL	<i>ecioka(y) ~ ecokay</i> ; W also <i>esioka(y)</i> ; <i>eciokay-ahcin</i>	Chiri (1942: §90); Murasaki (2025: 11)
3SG	<i>ani(hi)</i>	Chiri (1942: §90); Murasaki (1979: 146)
3PL	(<i>okay</i> , reported); bare <i>utara</i> ; <i>taranoka aynu utah</i>	Chiri (1942: §90); Sakaguchi (2020a: 189); Murasaki (2025: 11)

Table 19: Personal pronouns of Sakhalin Ainu

The pronouns are built on the person prefixes plus *ani* 'being' in the singular and *okay* 'exist.PL' in the plural; their grammatical number follows the series, not the reference, so *ciokay* and *anokay* are formally plural even when they mean 'I' (Chapter 13). Genre and dialect sort the first persons sharply. In our counts, the three

Sentoku Tarōji (千徳太郎治) letters use *cokay* 32 times, *anokay* 7 times, *eani* ~ *iani* 23 times, and *kuani* never; Asai Take's west-coast tales have *kuani* 38 times, *eani* 10, *anoka* 6, and *cokay* never. Chiri's consultants felt *kuani* to be the politer choice beside homely *anokay*, and treated *ciokay* and *anokay* as interchangeable (Chiri 1942: §90); the CI-pronoun is at home on the east coast and the southern west coast and marginal in Rayciska (Sakaguchi 2024: 27–29).

Syntactically the pronoun is an optional adjunct that doubles an argument already indexed on the verb, whatever its role: subject, object, or possessor (*kuani ku-cipihi ne* ‘it is MY boat’, (Chiri 1942: §91)):

- (1) *iineahsuy kuani hanka en-rayki*
 hey 1SG NEG 1SG.O-kill

‘Please — don’t kill ME!’ — Murasaki 2001: text 25; West Sakhalin, Odasu (Asai Take)
kuani doubles the object prefix *en-*.

- (2) *eani e-ye anpe neanpe neyra ne anahkayki hane an-nuu kusuiiki*
 2SG 2SG.A-say thing TOP whatever COP although NEG INDEF.A-hear going.to

‘Whatever YOU may say, we will not listen to it.’ — Sakaguchi 2020a: 191; East Sakhalin
 From Yamabe Yasunosuke's (山辺安之助) autobiography (Yamabe & Kindaichi 1913: 37); *eani* doubles *e-* inside a relative clause.

- (3) *cokay maci sine-p osura hemaka*
 1PL.EXCL wife one-NMLZ discard finish

‘I have put away one of my wives.’ — Tangiku & Ogihara 2001: Letter 1; East Sakhalin, Otasan (Sentoku Tarōji)
cokay with singular reference, as regularly in the letters.

- (4) *ecokay neanpe cis ohta eci-okay*
 2PL TOP boat in 2PL.S-exist.PL

‘As for you, you stay in the boat.’ — Pilsudski 1912: text 6, ll. 61–62; Sakaguchi 2020a: 197; East Sakhalin
 A curt command; Sakaguchi reads the bare pronoun as unceremonious beside the polite plural address with *utara* (Chapter 8).

9.1.1 Third person: *anihi* and its substitutes Third persons are normally bare nouns or nothing at all. The pronoun *ani(hi)* ‘he, she’ — formally the possessive of *an* ‘being’ — is rare (four tokens in the whole of Pilsudski's Materials, (Sakaguchi 2020a: 189)) and behaves unlike the speech-act pronouns: it can follow a noun phrase as an appositive intensifier, ‘X himself’ (Murasaki 1979: 78):

- (5) *tara aynu anihi naa oman kusu nee manu*
 that person 3SG too go.SG PURP COP REP

‘That man himself is going to go too, they say.’ — Murasaki 1979: 146; Sakaguchi 2020a: 198; West Sakhalin, Rayciska

- (6) *tani anahne anihi isam teh ohacirike ta*
 now TOP 3SG not.exist and absence LOC

‘Now that you are away, and in your absence...’ — Pilsudski 1912: text 21, ll. 190–192; Sakaguchi 2020a: 194; East Sakhalin

A wife to her husband: the third-person pronoun used of the hearer as respectful avoidance, the singular counterpart of honorific *utara* (Chapter 8).

For the plural, Chiri recorded *okay* ‘they’ from his consultants (Chiri 1942: §90), but no textual token has been found (Sakaguchi 2020a: 189). In the texts the slot is filled by bare *utara* (anaphoric ‘they’; Chapter 8) and, deictically, by the phrase *taranoka aynu utah* ‘those people’, which Murasaki gives as the conversational third person plural (Murasaki 2025: 11). We take Chiri's *okay* at face value as an elicitable but textually moribund form, squeezed out by *utara*.

9.2 The indefinite-person pronoun *anoka(y)*

anoka ‘I, we’ ~ *anokay* is the pronoun of the AN-series and shares its whole range — ‘we’, ‘I’, unspecified people, narrative first person — described in Chapter 13. Two points belong here. First, the dedicated plural *anokay-ahcin*, with the nominal plural =*ahcin* of Chapter 8, disambiguates ‘we’ from ‘I’; Hattori found it meaning ‘we’ for Rayciska speakers in their twenties and thirties but ‘I’ for men over sixty — the plural pronoun travelling

the same plural-to-singular road as the AN-series itself (Hattori 1961: 6), (Sakaguchi 2020b: 6). Yamabe's autobiography has *anokay-ahsin* 102 times against *anokay-utara* 14 (Sakaguchi 2020b: 6):

- (7) *yaykota* *e-wente* *yahka* *anokay-ahcin* *hene* *an-oskoro* *anpe* *kayki* *han* *ne*
oneself 2SG.A-break although INDEF-PL even INDEF.A-begrudge thing even NEG COP

‘Even if you break it yourself, it is not a thing that we would begrudge.’ — Pilsudski 1912: text 23, l. 201; Sakaguchi 2020a: 195; East Sakhalin

The brothers speak as a group: *anokay-ahcin* ‘we (and ours)’, with the AN-series prefix *an-* on the verb.

Second, like the other pronouns it is an emphatic adjunct, and it freely takes numeral appositions: *anoka tu-aynu-ne sap-an* ‘the two of us will go down’ (Chapter 8, example there). Within quoted dialogue it alternates with *kuani* by register, not by meaning; the choice of pronoun simply follows the person series chosen on the verb.

9.3 Demonstratives

Demonstratives divide into spatial (exophoric) and anaphoric series, and each form is either adnominal or pronominal, the pronominals being built from the adnominals with the nominalizers *-pe* ~ *-h* (Chiri 1942: §83, §97), (Murasaki 2025: 11, 13). Table 2 assembles the attested system.

	adnominal sg	adnominal pl	pronominal
proximal	<i>tan</i>	<i>tanoka(y)</i>	<i>tampe, tah</i>
distal	<i>tara, tara-an</i>	<i>taranoka(y)</i>	<i>tarampe, tarah</i>
far distal (emphatic)	<i>taan</i>	<i>taanoka(y)</i>	<i>taampe, taah</i>
anaphoric	<i>neya, nean</i> ~ E <i>nea(n)</i>	<i>neeroh, neokay</i>	<i>neampe, neyah</i>

Table 20: Demonstratives (W forms; E variants in the text)

The spatial system is two-term at base — speaker-near *tan* against speaker-far *tara*, with no ‘that (by you)’ : ‘that (yonder)’ split (Murasaki 2025: 11) — plus an emphatic far form. Hokkaidō distal *to(an)* is absent from Sakhalin, where *too* is only ‘lake’ and ‘day’; the distal slot is *tara(-an)*, literally ‘being there’, a Sakhalin–Kuril retention (Chiri 1942: §97). The plural adnominals contain *okay* ‘exist.PL’ and pluralise the noun phrase by themselves (*neeroh nuso-utara* ‘those sledges’, Chapter 8); a plural form is never used of a singular referent, though the singular may stand for plurals (Sakaguchi 2020b: 16). Only *tan* compounds with time nouns: *tanto* ‘today’, *tanpa* ‘this year’ (Chiri 1942: §99), (Murasaki 2025: 20).

- (8) *tara* *itanki* *nakorope* *hee*
that cup whose.thing Q

‘Whose is that cup?’ — Murasaki 2025: 16; West Sakhalin

The reply is *tan itanki ku-koro-pe hee* ‘this cup is mine’ — the minimal *tan* : *tara* pair.

- (9) *ee* *taranoka* *monimahpo* *uta* *nahta* *ampe* *eci-konopuruu*
well that.yonder girl PL which thing 2PL.A-like

‘Well then, you girls over there — which one do you like?’ — Murasaki 2001: text 17; West Sakhalin, Odasu (Asai Take)

9.3.1 Anaphoric *neya* and *nean* in narrative The anaphoric (“ideational”) series points not into space but into the discourse: *neya cise* ‘the aforementioned house’. Chiri stressed that Ainu, with no articles, uses exactly this series where English uses the definite article, and notes *nean* > *nan* from northern Hokkaidō into Sakhalin (Chiri 1942: §100–§101). The east-coast texts use *nea(n)*, reduplicable as *nea nea* ‘that same — you remember’ (Chapter 8, example (13) there). In the west-coast tales of Asai Take the anaphorics dominate the whole demonstrative economy. Our counts over the 3,852 corpus sentences: *neya* 437 tokens and *ne’an* 195 (with *neeroh* 7, *neyan* 10), against only 41 of proximal *tan*, 17 of *tara* and 7 of *taranoka* — the exophorics confined almost entirely to quoted dialogue, the anaphorics carrying reference-tracking in the narration. Add pronominal *tah* ‘this (just said)’ (198 tokens) and topicalised *ne’ampe* ‘that thing; as for that, thereupon’ (917 tokens, mostly grammaticalized as a topic-switching connective, Chapter 24), and the deictic word family makes up a remarkable share of running narrative text.

- (10) *neyke tani neya horokewpo utah okore taa asip-ahci manu*
and.then now that man PL all then go.out.PL-PL REP

‘And then those men all went out.’ — Murasaki 2001: text 37; West Sakhalin, Odasu (Asai Take)

- (11) *tani taa ahcahciso tah nuu kusu ene yee manu*
now then old.woman this hear because thus say REP

‘Now the old woman, having heard this, spoke thus.’ — Murasaki 2001: text 5; West Sakhalin, Odasu (Asai Take)
Pronominal *tah* resumes the preceding utterance; *ene* ‘thus’ points forward to the quotation.

9.3.2 Demonstrative adverbs Place adverbs are built on the same roots with the locative *ta*: *teeta* ‘here’, *taata* ‘there’ (Murasaki 2025: 16); Chiri adds nominal *teoro* ‘this place’ and *taani* ‘that place’ (Chiri 1942: §81, §83). The bare root *taa* is ubiquitous in Asai Take’s narration (some 6,700 tokens) as a discourse-deictic ‘there, then’, the rhythmic connective of spoken narrative; we analyse it as the medial root in adverbial function rather than as a filler without content, since it still contrasts with *teeta* in dialogue. Manner deictics are *ene* ‘thus, in this way’ (216 tokens), typically cataphoric before quotations as in example (11), and the quotative *nah* ‘so’ (Chapter 16); *tani* ‘now’ completes the set as the temporal proximal.

- (12) *taata hemata an-i hii*
there what exist.SG-EP Q

‘What is over there?’ — Murasaki 2025: 16; West Sakhalin

The answer is *teeta itanki an* ‘here is a cup’; *-i* is the epenthetic vowel before the final particle *hVV*.

9.4 Interrogatives

The interrogatives of Sakhalin Ainu are nouns and adverbs built on three roots, *hem-*, *ne-* ~ *na-* and *tema-* (Chiri 1942: §104). The attested inventory, with token counts from the Asai Take corpus where it occurs there: *hemata* ‘what’ ‘what’ (162), *naata* ‘who’ ‘who’ (6; also *nata* in Pilsudski’s texts, (Sakaguchi 2020a: 195)), *temana* ‘how’ ‘how, what kind of’ (22), *nahta* ‘where; which’ (2), *neyta* ‘where’ (1), *nakene* ‘whither’, *nahwa* ‘whence’, *nakoro(ke)he* ‘where; whose’, *hempara* ‘when’, *hempah(no)* ‘how many, how much’, *hemata kusu* ‘why’ (Murasaki 2025: 19). Adnominal interrogatives are *hemanu* ‘which’, connective-built *neran* and *temana-an* ‘what kind of’ (Chiri 1942: §104). Hokkaidō *nen* ‘who’ and *nep* ‘what’ are unattested in Sakhalin material — in Pilsudski *nen* occurs only as a variant of *nenō* ‘like’ (Pilsudski 1912: 247) — and the person interrogative is uniformly *na(a)ta*.

- (13) *hemata kamuy kam*
what god meat

‘The meat of what creature is this?’ — Murasaki 2001: text 30; West Sakhalin, Odasu (Asai Take)

The answer is *seta kam* ‘dog meat’; *kamuy* covers game animals.

- (14) *tara aynu naata*
that person who

‘Who is that person?’ — Murasaki 2025: 10; West Sakhalin

- (15) *temana e-kii teh e-makan-i hii*
how 2SG.S-do and 2SG.S-go.uphill.SG-EP Q

‘How did you come to climb up here?’ — Murasaki 2001: text 34; West Sakhalin, Odasu (Asai Take)

Interrogatives stand in situ in the position of the questioned constituent, usually with the final particle *hii* ~ *hee* closing the clause (Murasaki 2025: 18); there is no interrogative inversion. In content questions the verb commonly appears in its nominalized form (*e-makan-i* above; (Chiri 1942: §40)).

9.5 Indefinite uses: interrogative + *ka*

As across the family, interrogatives double as indefinites when followed by *ka* ‘even, also’: *hemata ka* ‘something’ (nine tokens in the Asai corpus), *temana ka* ‘somehow, in some way or other’ (five), *neyta ka* ‘somewhere’, *hempara ka* ‘at some time, ever’ (Murasaki 2025: 51). Under negation the same string is a negative indefinite, ‘nothing, nobody’. Beside the interrogative roots the defective noun *nee* ~ *neh* ‘some(thing)’ serves the same construction: *nee aynu ka isam* ‘there is nobody’, *neera ampe ka isam* ‘there is nothing’ (Murasaki

2025: 17), and Dobrotvorskyy's nineteenth-century conversations already show *nera ham ariki kusuiki* 'no one (of any kind) will come' (Dobrotvorskyy 1875), (Sakaguchi 2020a: 190).

- (16) *hemata ku-yee hee neh ka ku-yee ka hankii*
 what 1SG.A-say Q some even 1SG.A-say even NEG

'What did I say? I said nothing at all.' — Murasaki 2001: text 36; West Sakhalin, Odasu (Asai Take)

- (17) *sine cise an-i ike taa ohta nee aynu ka isam manu*
 one house exist.SG-NMLZ and then there some person even not.exist REP

'There was a house, and there was no one at all in it.' — Murasaki 2001: text 8; West Sakhalin, Odasu (Asai Take)
 sine 'one, a certain' introduces the new referent (Chiri 1942: §100); nee ... ka ... isam is the regular 'nobody'.

9.6 Reflexives and intensifiers

There are no reflexive or reciprocal pronouns: the verbal prefixes *yay-* ~ *si-* 'self' and *u-* 'each other' do that work (Chiri 1942: §81) (Chapter 14). What the language does have are adverbial intensifiers: *yaykota* 'oneself, of one's own accord', attested in Pilsudski's texts (*yaykota e-koro kun-pe ne* 'it is something you must keep yourself', and example (7) above; (Sakaguchi 2020a: 195)); *sinenehka* 'alone, by oneself', with 35 tokens the everyday form in Asai Take's tales; and appositive *anihi* 'X himself' (§9.1.1). The emphatic doubling function that other languages assign to reflexive pronouns falls in Sakhalin Ainu to the personal pronouns themselves, which exist for little else.

Chapter 10 · Numerals and quantification

The vigesimal numeral system, classifier suffixes, ordinals, and quantifiers.

Sakhalin Ainu numerals fall into two formally distinct series: a bare *adnominal* series (*sine* 'one', *tu* 'two', *re* 'three'...), which can only stand before a noun, and a *nominal* series derived from it (*sineh* 'one', *tuh* 'two', *reh* 'three'...), which is used for counting and stands in the clause like any other noun. Chiri Mashiho 知里真志保 built his classic account of Ainu numerals around exactly this opposition Chiri (1942: §§105–113), and the Sakhalin corpus bears it out throughout. On top of the inherited base-twenty arithmetic, the Sakhalin variety developed a productive decimal system around the borrowed base words *kunkutu* 'ten' and *tanku* 'hundred' — the fullest studies are Murasaki (Murasaki 2009) and Sakaguchi (Sakaguchi 2022). This chapter sets out the cardinal numerals (§10.1), the vigesimal system and the decimal innovation (§10.2), the counting of humans and other unit-noun constructions (§10.3), the general quantifiers (§10.4), numeral syntax and calculation practice (§10.5), and the expression of ordinal rank (§10.6).

10.1 Cardinal numerals one to ten

The basic numerals are adnominals: they modify a following noun and never stand alone as an argument Chiri (1942: §105). To count objects, or to use a numeral as a noun phrase by itself, the adnominal is nominalized with the 'thing' suffix that appears as *-h* after vowels and *-pe* after consonants; the *-h* of *sineh*, *tuh* corresponds to Hokkaidō *-p* by the regular coda neutralisation described in Chapter 4 Chiri (1942: §113), (Murasaki 2009: 74). Table 1 gives both series.

	adnominal	nominal	sources
1	<i>sine</i>	<i>sineh</i>	(Murasaki 2025: 24–25); (Chiri 1942: §106); both pervasive in the corpus, e.g. (1), (2)
2	<i>tu</i>	<i>tuh</i>	(Murasaki 2025: 24–25); corpus <i>tuh pis</i> , see (3)
3	<i>re</i>	<i>reh</i>	(Murasaki 2025: 24–25); corpus <i>re horokewpo</i> ‘three men’ (Murasaki 2001: text 51)
4	<i>iine</i>	<i>iineh</i>	(Murasaki 2025: 24–25); (Chiri 1942: §106) (<i>ine</i>)
5	<i>asne ~ asisne</i>	<i>asneh ~ asisneh</i>	(Murasaki 2025: 24–25); <i>asisne</i> in the east and in Yamada Hayo's speech, cf. (12), <i>asisneh</i> in Dobrotvorsky (Sakaguchi 2022: 69)
6	<i>iwan</i>	<i>iwampe</i>	(Murasaki 2025: 24–25); corpus <i>iwan monimahpo</i> ‘six girls’ (Murasaki 2001: text 51)
7	<i>arawan ~ aruwan</i>	<i>arawampe</i>	(Murasaki 2025: 24–25); <i>aruwan</i> in Yamabe's text, cf. (Sakaguchi 2022: 73)
8	<i>tupesan</i>	<i>tupesampe</i>	(Murasaki 2025: 24–25); (Chiri 1942: §112)
9	<i>sinepisan ~ sinepesan</i>	<i>sinepisanpe</i>	(Murasaki 2025: 24–25); <i>sinepesan</i> (Chiri 1942: §112)
10	<i>wan</i>	<i>wampe</i>	(Murasaki 2025: 24–25); (Chiri 1942: §106)

Table 21: The cardinal numerals 1–10. The Rayciska forms are those of Murasaki's course book and ultimately of Hattori's survey; divergent forms from other localities are noted. The adnominal series is identical across the island except for 4, 5, 7, and 9, where localities diverge.

Murasaki records the nominal series as the counting chant itself: *sineh*, *tuh*, *reh*, *iineh*, *asneh*, *iwampe*, *arawampe*, *tupesampe*, *sinepisanpe*, *wampe* Murasaki (2025: 24). Two points of variation cut across the table. First, ‘five’: the west-coast Rayciska form *asneh* has simplified the cluster of expected **asisneh*, while Dobrotvorsky's mid-nineteenth-century west-coast records and east-Sakhalin speakers keep *asisne(h)* Murasaki (2009: 74), (Sakaguchi 2022: 69); both shapes occur in the corpus, cf. (12). Second, the final consonant: Naert's informants from Taraytomari and Nairo pronounced final *-p* (*sinep*), so the *-h* shapes of Table 1 must not be projected onto every locality Sakaguchi (2022: 69).

The two series divide the labour cleanly in texts. The adnominal counts a following noun:

- (1) *sine* *acahcipo* *taa* *mici-hi-hcin* *tura* *an* *manu*
 one old.woman EMPH grandchild-POSS-PL COMIT exist.SG REP

‘An old woman lived there with her grandsons.’ — Murasaki 2001: text 51; West Sakhalin, Odasu (Asai Take)

The nominal form stands alone, as subject, object, or — very commonly — followed by the quantifier particles of Chapter 24:

- (2) *neeteh* *tani* *taa* *sineh* *pateh* *taa* *an* *manu*
 and.then now EMPH one only EMPH exist.SG REP

‘And then only one (of the dogs) was left.’ — Murasaki 2001: text 37; West Sakhalin, Odasu (Asai Take)

With ‘two’ and ‘three’ the West-coast narratives very often add the unit noun *pis* ‘piece’ — *tuh pis*, *reh pis* — whether the referents are objects or people; the combination is the default way Asai Take counts pairs (24 tokens of *tuh pis* in her tales). Chiri records the same element for Hokkaidō as an optional extension *tu-p(-pis)* (Chiri 1942: §113).

- (3) *neya* *monimahpo* *tuh* *pis* *taa* *ehahtaa-hci*
 that girl two piece EMPH dig.lily.roots-PL

‘Those two girls went digging lily roots.’ — Murasaki 2001: text 19; West Sakhalin, Odasu (Asai Take)

tuh pis in apposition to the subject; for another example see (11) of Chapter 6.

sine doubles as an indefinite article-like determiner ‘a certain, one’ Chiri (1942: §115): story openings of the type *sine horokewpo an manu* ‘there was a man’ and the temporal formula *sine too* ‘one day’ (over a hundred

tokens in the Asai corpus) are the normal way a new participant or scene is introduced. Beside *sine*, the old proclitic *ar-* ‘one (of a pair)’ survives in fixed formations: ‘once’ is *ar-suy*, never **sine suy* Chiri (1942: §113).

10.2 The vigesimal system and the decimal innovation

From eleven to nineteen every recorded Sakhalin variety, from Dobrotvorský's dictionary to Hattori's Rayciska survey, uses one and the same additive pattern: the unit numeral comes first, then *ikasma* ‘to be left over’, then *wampe* ‘ten’ — *sineh ikasma wampe* ‘eleven’, *tuh ikasma wampe* ‘twelve’, literally ‘one left over, (and) ten’ Chiri (1942: §107), Sakaguchi (2022: 70), (Murasaki 2025: 25). The element *ikasma* is the ordinary verb ‘be left over, remain’, alive in the letters of Sentoku Tarōji 千徳太郎治:

- (4) *kara ikasma ciki ihoku yahka pirika manuy*
make be.left.over if sell though be.good REP

‘If any are left over after making (them), they say it is fine to sell them.’ — Tangiku & Ogihara 2001: Letter 1; East Sakhalin, Otasan (Sentoku Tarōji)

10.2.1 The score *hoh* ‘Twenty’ is *hoh* ‘twenty’ (Hokkaidō *hot*), and unlike the numerals of Table 1 it is a noun; its adnominal is derived with *-ne*, *hohne* Chiri (1942: §108). The inherited system counts in scores: *tu hoh* ‘forty’, *asisne hoh* ‘hundred’, *iwan hoh* ‘hundred and twenty’ Chiri (1942: §109). The intervening tens are ‘overcounted’ toward the next score with the applicative *e-*: *wan e-tu-hoh* ‘thirty’, literally ‘ten (going) to forty’ Chiri (1942: §110), and units are undercounted with *ikasma*, *asisne ikasma hohne* ‘twenty-five’ (Chiri 1942: §111).

The two devices are not in free variation: they divide each score at its midpoint. Below it, the count hangs from the score just passed — twenty-one to twenty-nine are *ikasma*-forms on *hohne*; above it, the count reaches for the score ahead — thirty-one is *sine ikasma wan e-tu-hoh*, ‘(one over ten) toward forty’. Ochiai, who states the rule in these terms, observes that Hokkaidō and Sakhalin agree on the switch point in detail and therefore reconstructs the undercounting-to-overcounting alternation itself for late Proto-Ainu, rather than treating the *e-* forms as a local elaboration; her closest typological parallels are the overcounted vigesimal systems of Tibeto-Burman languages Ochiai (2021).

This machinery was fully productive in the west-coast speech that Dobrotvorský recorded around Kusunnay in 1867–72 (Dobrotvorský 1875), and there it scales: *hoh* serves as base only up to 199, and each higher base is the tenfold of the one below it Sakaguchi (2022: 78–80).

value	base	formation	serves as base for
20	<i>hoh</i>	inherited score	20–199
200	<i>sinewano</i>	fused <i>sine-wan-hoh</i> ‘one ten of scores’	200–1,999
2,000	<i>waysinewano</i>	‘ten two-hundreds’	2,000 upward
4,000,000	<i>waysinewano waywanho</i>	2,000 × 2,000, multiplicative	ceiling of the record

Table 22: The scaling bases of Dobrotvorský's vigesimal system (Kusunnay, 1867–72), as reconstructed by Sakaguchi (2022: 78–80). Each base is ten times the previous one, so the architecture is vigesimal in its bricks but decimal in its storeys; 400 and 8,000, the doubled values a pure score system would call for, are accordingly not bases at all.

At the top the dictionary simply multiplies the largest base into itself — *waysinewano waywanho* ‘four million’, a figure far beyond anything recorded for Hokkaidō Sakaguchi (2022: 79–80). Below the bases, units are slotted in with *ikasma* at each level:

- (5) *sineh ikasma sinehoh ikasma sinewano*
one be.left.over one.score be.left.over two.hundred

‘two hundred and twenty-one’ — Dobrotvorský 1875: 305; Sakaguchi 2022: 79; West Sakhalin, Kusunnay area
Dobrotvorský writes ⟨sine⟩; the final -h is supplied by Sakaguchi.

10.2.2 The decimal bases *kunkutu*, *tanku*, *wantanku* Alongside this, every source from Dobrotvorský onward attests a decimal system built on *kunkutu* ‘ten’ and *tanku* ‘hundred’, which are nouns usable directly as numerals. Their origin is partly clear: *tanku* matches Manchu–Tungusic words for ‘hundred’ (Manchu

tanggū) and Pilsudski was told it came from the ‘Olcha’ with the marten-trapping trade Pilsudski (1912: 141), (Murasaki 2009: 75), while *kunkutu* has no good Tungusic match and may be an internal Ainu word pressed into service for trade counting Sakaguchi (2022: 66). By the twentieth century the decimal forms had displaced the scores everywhere: Naert’s mid-century informants from Taraytomari and Nairo knew only *kunkutu* forms for ‘thirty’, and his Rayciska informant offered *wampe ikasma sine hoh* only as an archaism Sakaguchi (2022: 71). Hattori’s dictionary likewise lists *hohne* ‘twenty’ as ‘elders’ speech’ (*onne itah*) Hattori (1964: 262), (Murasaki 2009: 74).

	form	sources
10	<i>wampe</i>	(Murasaki 2025: 25); (Hattori 1964: 262); never <i>*sine kunkutu</i> (Murasaki 2009: 74)
20	<i>tu kunkutu</i> (elders <i>hohne</i>)	(Murasaki 2025: 25); (Hattori 1964: 262)
30	<i>re kunkutu</i>	(Murasaki 2025: 25); (Sakaguchi 2022: 71)
50	<i>asne kunkutu</i>	(Murasaki 2025: 25)
100	<i>sine tanku</i>	(Murasaki 2025: 25); (Chiri 1942: §112)
110	<i>sine tanku orowa wampe</i>	(Hattori 1964: 264); (Sakaguchi 2022: 76)
200	<i>tu tanku</i>	(Murasaki 2025: 25)
1,000	<i>sine wantanku</i>	(Murasaki 2025: 25); (Sakaguchi 2022: 72–73)
2,000	<i>tu wantanku</i>	(Murasaki 2025: 25); also from an Odasan speaker in 1967 (Sakaguchi 2022: 73)

Table 23: The decimal series (Rayciska), after Hattori’s survey as tabulated by Murasaki; the same forms recur in Laufer’s 1899 east-coast list and Kindaichi’s 1907 records as reviewed by Sakaguchi.

Note the asymmetry preserved at ‘ten’: the bare decade is the native *wampe*, never **sine kunkutu*; the borrowed base begins at twenty Murasaki (2009: 74), (Sakaguchi 2022: 71–72). ‘Thousand’ is the transparent compound *wantanku* ‘ten hundreds’, so Murasaki could elicit the year 2008 as *tu wantanku orowa tupesan paa* ‘two thousand and then eight years’ (Murasaki 2009: 75).

The decimal system in running text is best observed in the autobiography that Yamabe Yasunosuke 山辺安之助 dictated in 1912 (あいぬ物語), analysed in detail by Sakaguchi. Score forms do occur there, mostly with years and days:

- (6) *meiji iwan paa ikasma hohne paa nee paakehe*
Meiji six year be.left.over twenty year COP year.POSS

‘the year Meiji 26 (= 1893)’ — Sakaguchi 2022: 73; East Sakhalin, Yamabeci (Yamabe Yasunosuke)

From Yamabe’s *あいぬ物語* (1913: 5); *hohne* too ‘twenty days’ recurs there twice. Only six *hoh*-forms occur in the whole book, against sixteen with *tu kunkutu*.

But the workaday counting is decimal, and the same value can be assembled in two ways — by bare juxtaposition on the Sino-Japanese model, or with native *ikasma*:

- (7) *tu kunkuto asisne aynu*
two ten five person

‘twenty-five men’ — Sakaguchi 2022: 75; East Sakhalin, Yamabeci (Yamabe Yasunosuke)

あいぬ物語 p. 114.

- (8) *tu kunkutu ikasma asisne aynu*
two ten be.left.over five person

‘twenty-five men’ — Sakaguchi 2022: 75; East Sakhalin, Yamabeci (Yamabe Yasunosuke)

あいぬ物語 p. 18 — the same value as (7), built with *ikasma*; Murasaki (1979: 88) records the mirror order *sineh ikasma tu kunkutu* ‘21’ from Rayciska, so the decade constructions wavered even within one speaker’s usage.

- (9) *tupesan tanku asisne kunkutu aynu*
eight hundred five ten person

‘eight hundred and fifty people’ — Chiri 1942: §112; Sakaguchi 2022: 77; East Sakhalin, Yamabeci (Yamabe Yasunosuke)

Large sums show the two systems interleaved. In Chiri's example of current usage the hundreds are counted by a score:

- (10) *tu ikasma hoh tanku sinepesan kunkutu iine komaska*
two be.left.over twenty hundred nine ten four money

‘2,294 yen’ — Chiri 1942: §112; Sakhalin (dialect not further specified)

The first three words form a unit: [tu ikasma hoh] ‘22’ multiplies tanku — ‘22 hundreds, 9 tens, 4’ (Sakaguchi 2022: 80–81).

No base for ‘ten thousand’ is recorded anywhere (Hattori 1964: 265), (Sakaguchi 2022: 64); beyond a thousand, speakers either stretch the multiplier — adnominal *hohne* ‘twenty-fold’ on *wantanku* — or juxtapose number nouns as an outright multiplication:

- (11) *hohne wantanku re kunkutu wantanku pahno an ikoro*
twenty.times thousand three ten thousand as.much.as exist.SG money

‘money amounting to twenty or thirty thousand yen’ — Sakaguchi 2022: 74; East Sakhalin, Yamabeci (Yamabe Yasunosuke)

hohne wantanku ‘20×1000’ uses the adnominal of *hoh* as a coefficient; *re kunkutu wantanku* ‘30×1000’ juxtaposes the number noun ‘thirty’ with ‘thousand’ — Sakaguchi’s ‘multiplicative’ type.

10.3 Counting humans and other unit nouns

Hokkaidō Ainu counts humans with a dedicated series in *-n* (after vowels) and *-iw* (after consonants): *sinen*, *tun*, *iwaniw* Chiri (1942: §113). For Sakhalin, Murasaki states that this series is absent and that humans are counted with adnominal numeral plus *aynu* ‘person’ — *sine aynu* ‘one person’, *iwaniw aynu* ‘six people’ Murasaki (2009: 75), (Murasaki 2025: 25). The corpus does confirm the *aynu* pattern as the productive one:

- (12) *asisne aynu ka iwan aynu ka an*
five person even six person even exist.SG

‘There were as many as five or six people.’ — Kitahara et al. 2003; Sakhalin (dialect not further specified), Yamada Hayo

The statement of absence does not survive the record, however. Sentoku's east-coast letters use *sinen* in a negative-polarity idiom:

- (13) *sinen ka isan*
one.person even not.exist

‘There is not even a single person (left).’ — Tangiku & Ogihara 2001: Letter 2; East Sakhalin, Otasan (Sentoku Tarōji)

More decisively, Chiri — whose grammar describes Sakhalin, not Hokkaidō — sets the bound series out as a complete paradigm: *sinen*, *tun*, *ren*, *inen*, *asiknen* with *-n* after vowels, then *iwaniw*, *arawaniw*, *tupesaniw*, *sinepesaniw*, *waniw* with *-iw* after consonants, closed by the interrogative *hempakiw* ‘how many people’ Chiri (1942: §113). These are his stated Sakhalin forms, and they do duty in his plain prose examples:

- (14) *menoko sine-n ek*
woman one-CL.HUM come

‘One woman came.’ — Chiri 1942: §113; Sakhalin (dialect not further specified)

The bound counter even supplies the unit half of the additive teen frame, where the counted noun is named only once:

- (15) *tu-n ikasma wan ci-yupi*
two-CL.HUM be.left.over ten 1PL-elder.brother

‘my twelve elder brothers’ — Chiri 1942: §114; Sakhalin (dialect not further specified)

The unit slot holds the bound form *tun*, not *tu aynu*; the head noun *yupi* appears once, on the ten.

The narrative corpus agrees with Chiri rather than with the claim of absence. In Asai Take's tale of the boat-builders, the flying chips are counted off in a chant *sinen kuh ray*, *tunne kuh ray*, *renne kuh ray* ‘one man dead, two men dead, three men dead’ (Murasaki 2001: text 32), preserving *sinen* and, we take it, *tun* and *ren* extended by *-ne*; and Piłsudski's east-coast corpus has *iwaniw-niw* ‘six (rowers)’ in a clause that also shows *aynu enko* ‘half of the men’ with *enko* ‘half’ Piłsudski (1998: 287), (Dal Corso 2024: 75). Sakhalin therefore kept *both* strategies side by side: the inherited, phonologically conditioned *-n/-iw* counter — complete in Chiri, persisting in chants, formulae, and the low numerals — and the transparent innovation with *aynu*, which had become the default by the last-recorded generation. Murasaki's ‘absent’ captures the end state of that competition among her

Rayciska informants, not a gap in the variety; the pairing of an opaque inherited counter with a semantically transparent renewal is a familiar one cross-linguistically, and here the renewal won.

Other unit nouns combine with the adnominal series in the same way as *pis* and *aynu*. Days: *sine too* ‘one day’, *iwan too* ‘six days’, with suppletive *tutko*, *rerko* ‘two, three days’ (Chiri 1942: §113). Years and months: *paa* and *cuh*, as in (6) and in Sentoku’s *sine cup okta* ‘per month’ (Tangiku & Ogihara 2001: Letter 1). Times: *suy*, with *ar-suy* ‘once’ (§10.1). Fathoms: *tem*, in the corpus fused in *iwanten kosonto* ‘a six-fathom robe’ (Murasaki 2001: text 51). Chiri further records a Sakhalin-only collective *atuyta* ‘ten head of game’ for counting catch (Chiri 1942: §113).

10.4 Quantifiers

The general quantifiers are adverb-like words that stand before the verb or before the noun phrase they measure. ‘Many, in numbers’ is *poronno* ‘many’ (from *poro* ‘big’) and, in the West-coast narratives, above all *renkayne* ‘plenty’, which in Asai Take’s tales is the routine word for ‘in great numbers’ (64 corpus tokens) and which Murasaki’s course book likewise uses for ‘a lot’ (Murasaki 2025: 22); its other, clause-linking life (‘because of’) is taken up in Chapter 11.

- (16) *tek-ih-i-cin* *ohta* *neya* *ay* *ne* *kun-pe* *poronno* *uyna*
hand-POSS-PL LOC that arrow COP should-NMLZ many take.PL

‘In his hands he grabbed up many of the things that were to be those arrows.’ — Pilsudski 1912: 70; Chiri 1942: §67; East Sakhalin

- (17) *neya* *kamuy* *utah* *renkayne* *wahka* *ohta* *hecire-hci* *wa* *okay-ahci* *manu*
that god PL plenty water LOC dance-PL and exist.PL-PL REP

‘Those gods were dancing in crowds in the water.’ — Murasaki 2001: text 46; West Sakhalin, Odasu (Asai Take)

‘A little’ is *ponno* ‘a little’ (from *pon* ‘small’), often reduplicated *ponno ponno* ‘just a little’ (Murasaki 2001: text 39):

- (18) *naa* *ponno* *naycaata* *sanke* *kusu*
more a.little riverbank-LOC move.shoreward PURP

‘I will bring it a little closer to the bank.’ — Murasaki 2001: text 25; West Sakhalin, Odasu (Asai Take)

‘All’ is expressed by *okore* ‘all’ — the high-frequency exhaustive of the West-coast corpus, ‘all of them, completely’ (Chapter 8) — and by *emuyke* ‘all’, which quantifies over a preceding noun phrase:

- (19) *kawre* *nii* *emuyke* *ci-kehke* *wa* ...
wither tree all RES-break and ...

‘the withered trees all snapped, and ...’ — Kitahara et al. 2003; Sakhalin (dialect not further specified), Yamada Hayo

The Hokkaidō *shape opitta* ‘all’ is absent from the Sakhalin corpus, though its regular reflex *opista* (Chapter 4 (Phonology)) does appear in the Sentoku letters (Chapter 14 (Valency-changing morphology)); *okore* and *emuyke* are the usual exponents. Chiri cites *opitta* only from Hokkaidō sources, distinguishing it from *epitta* ‘all over’ (Chiri 1942: §123(104), (131)). Negative quantification uses *ka* ‘even’ with the negative existential, as in (13) and *nee aynu ka isam* ‘there is nobody’ (Murasaki 2025: 17); see Chapter 17. The interrogative quantifier is *hempah* ‘how many’ (adverbial *hempahno* ‘how much’, (Murasaki 2025: 24)), which with following *ka* yields an indefinite plural: *hempah too ka* ‘for some days’, *hempah poo ka korohci* ‘they had several children’ (Murasaki 2001: texts 48, 12).

10.5 Numeral syntax and calculation

Within the noun phrase the adnominal numeral immediately precedes its noun (*re horokewpo* ‘three men’), like any other adnominal of Chapter 6. When a counted noun phrase is built additively, the classical pattern repeats the noun or replaces its first occurrence with a nominal numeral: *sine-p ikasma wan suma* ‘eleven stones’, literally ‘one (thing) left over, ten stones’ Chiri (1942: §114). With the borrowed bases this constraint lapses — *kunkutu* and *tanku* are themselves nouns, so the counted noun follows the whole numeral phrase without being repeated, as in (7)–(9) Sakaguchi (2022: 75–76).

Used without a noun, the nominal numerals function as floating quantifiers in apposition to an argument, as in (3) and (14); the distributive particle *ranke* ‘each’ (Chiri 1942: §125(172)) combines with them:

- (20) *re* *horokewpo* *sineh* *ranke* *sam*
 three man one each marry

‘The three men each took one (of the girls) as wife.’ — Murasaki 2001: text 51; West Sakhalin, Odasu (Asai Take)

As for calculation practice, Sakaguchi’s survey shows a system that could be operated at scale. Addition within a base is handled by *ikasma* and, between bases, by *orowa* ‘and then’ — *sine tanku orowa sineh ikasma wampe* ‘111’ (Hattori 1964: 264), (Sakaguchi 2022: 76): *orowa* marks the point where the otherwise descending sequence of values gives way to the fixed ascending unit-*ikasma*-ten block. Multiplication is expressed either by an adnominal coefficient (*hohne wantanku* ‘20,000’) or by juxtaposing number nouns (*re kunkutu wantanku* ‘30,000’, (11)); Dobrotvorskij’s dictionary runs the same devices up to *waysinewano waywanho* ‘4,000,000’ Sakaguchi (2022: 79–80). The trade context that drove this — counting marten pelts and cash — is visible in the dictionary’s own glosses, where *kunkutu* and *tanku* first appear as tallies of marten snares Sakaguchi (2022: 66), and in the money counting of Murasaki’s conversation course, *sine wantanku en* ‘a thousand yen’ (Murasaki 2025: 24).

10.6 Ordinals

There is no ordinal affix. The Hokkaidō pattern with *ikin ne* (*iwan ikin ne* ‘sixth’) has no Sakhalin counterpart: the cognate *ikinne* survives only in its older adverbial sense ‘in a row, all together’ (*sine ikinne* ‘all as one’), never as an ordinal former (Chiri 1942). The previous descriptions give no Sakhalin ordinal paradigm, and the texts show why none was there to elicit: ranking is handled by two lexical items at the salient low end and by nothing at all above them. This section is our generalisation from the corpus.

‘First’ is suppletive. It is carried by the positional *hoski* ‘first’ ‘first, earlier’ (adverbial *hoskino*), with no formal connection to *sine* ‘one’ — the same dissociation of ‘first’ from ‘one’ seen in English *first* or Latin *primus*. Its standing narrative opposite is *yooponi* ‘later, after’:

- (21) *hoski* *tah-ci* *nii* *ka* *an* *yooponi* *tah-ci-pe* *ka* *an*
 first cut-PL tree even exist.SG later cut-PL-NMLZ even exist.SG

‘There was wood they had cut earlier, and wood they had cut later.’ — Murasaki 2001: text 34; West Sakhalin, Odasu (Asai Take)

‘Second, the next’ is the single derived form, (*i*)*otutanu*, built on *tu* ‘two’ (cf. Hokkaidō (*o*)*tutanu*); in the Asai corpus it names the second-eldest of a set of siblings. For sets, the lexical series *kiyanne* ‘eldest’, *inoskun* ‘middle’, *poniwnne* ‘youngest’ does the same work — the standard framing of the three-brother tales (Murasaki 2001: text 5).

From ‘third’ on, the bare adnominal cardinal stands before the noun and the ordinal reading comes from the narrative sequence alone. In Asai’s tale the hero works through the corners of a house one by one, and the third is simply *re* ‘three’:

- (22) *re* *cise-he* *sikhew* *kaari* *iciw*
 three house-POSS corner via stab

‘He stabbed through the third corner of the house.’ — Murasaki 2001: text 92; West Sakhalin, Odasu (Asai Take)
 Bare adnominal *re* read as ‘third’ from the running count of corners in the tale.

The Sakhalin ordinal system is thus suppletive at the bottom and zero-marked above: *hoski*(*no*) ‘first’, (*i*)*otutanu* ‘second, next’, and thereafter the plain cardinal disambiguated by the order in which the referents were introduced — typologically the commonest of ordinal systems, since ‘first’ (and often ‘second’) are cross-linguistically the positions most prone to suppletion while the rest of the series goes unmarked.

Chapter 11 · Postpositions and relational nouns

Case postpositions, relational (locative) nouns, and adverbial particles.

Sakhalin Ainu marks the role of oblique noun phrases with postposed case particles: *ohta* ‘in, at’ for location, *onne* ‘to, toward’ for direction, *orowa* ‘from’ and *wa* ‘from’ for source, *ani* ‘with, by means of’ for instrument, *tura* ‘together with’ for accompaniment, *ne* ‘as, into’ for a resulting or assumed state, and a handful of others. Chiri’s grammar classifies these as the ‘fourth class of particles’ (格助詞) and itemises them with their Sakhalin variants Chiri (1942: §§122–123); Murasaki’s course book presents the same core set for Rayciska (Murasaki 2025: 13, 21–22). Finer spatial relations are expressed not by particles but by *relational nouns* — *kaske* ‘top’, *empoke* ‘underside’, *onnayke* ‘inside’, *sam* ‘beside’ — which head a possessive-like phrase and then combine with the case particles (Chiri’s ‘first-class formal nouns’, (Chiri 1942: §§78–79); cf. Chapter 7). This chapter describes the particles (§11.1–§11.7), the relational nouns (§11.8), and the fuzzy boundary between case particle, relational noun, and clause linker (§11.9).

11.1 The case-particle system

Table 1 surveys the system. The particles follow the noun phrase directly; the noun keeps its conceptual or possessed form as the sense requires, and plural *-hcin* precedes the particle (*tek-ih-cin ohta* ‘in his hands’, (Pilsudski 1912: 70); (Chiri 1942: §67)). Most of the particles also have an independent, anaphoric use without a host noun — *ohta* ‘there’, *orowa* ‘then, after that’ (944 corpus tokens, mostly as a narrative connective), *turano* ‘at the same time’ — a point Chiri makes for each entry Chiri (1942: §123).

form	function	sources
<i>ohta</i> (<i>okta</i>), <i>ta</i>	locative–dative ‘in, at, to’	(Chiri 1942: §123(139–140), (150)); (Murasaki 2025: 21)
<i>onne</i>	allative ‘to, toward, into’	(Chiri 1942: §123(129)); (Murasaki 2025: 21)
<i>ene</i>	allative ‘to’, chiefly on relational nouns	(Chiri 1942: §123(103)); corpus, see (18)
<i>wa</i> , <i>orowa</i> , <i>oro</i>	ablative ‘from’; agent of a passive	(Chiri 1942: §123(157), (137), (135)); (Murasaki 2025: 21)
<i>ani</i>	instrumental ‘with, by means of’; cause	(Chiri 1942: §123(98)); (Murasaki 2025: 13)
<i>tura</i> (<i>no</i>)	comitative ‘together with’	(Chiri 1942: §123(153–154))
<i>ne</i>	mutative–essive ‘as, into’	(Chiri 1942: §123(121))
<i>pahno</i> (<i>pak no</i>)	terminative ‘until, as far as, as much as’	(Chiri 1942: §123(142)); (Murasaki 2025: 22)
<i>kaari</i> , <i>okakara</i>	prosecutive ‘through, along, via’	(Chiri 1942: §123(109), (127)); corpus, see (11)
<i>peka</i>	perlative ‘over, about (an area)’	(Chiri 1942: §123(141))
<i>oykari</i>	‘around’	(Chiri 1942: §123(126))
<i>un</i>	attributive ‘of, belonging to (a place)’	(Chiri 1942: §123(156))
<i>nah</i>	quotative	(Chiri 1942: §123(120)); Chapter 22

Table 24: Case particles of Sakhalin Ainu. Item numbers are Chiri’s entries under his §123; forms in parentheses are East-coast spellings in the Sentoku letters.

Two members of the Hokkaidō system are conspicuously restructured here. The allative is *onne*, a Sakhalin form Chiri flags as such (Hokkaidō uses *un* and *orun*); and the plain ablative is *wa* with its expansions *oro wa* > *orowa* and, in Sakhalin only, bare *oro* Chiri (1942: §123(129), (135), (157)). The attributive *un* survives in place-name epithets of the type *Sannupis-un re horokewpo* ‘the three young men of Sannupista’ (Murasaki 2001: text 51).

11.2 Locative and dative *ohta*, *ta*

ohta is the workhorse of the system — over five hundred tokens in the Asai corpus alone. Historically it is the relational noun *oro* ‘place’ plus the locative particle *ta* (*or-ta*), and Chiri notes that the fusion had gone so far that *ohta*, *orun*, *orowa* behave as simple particles Chiri (1942: §79). In the East-coast letters it is spelled *okta*, with the cluster unlenited (Tangiku & Ogihara 2001: Letter 1); see Chapter 5 on the *h ~ k* alternation. It marks both static location (1) and the goal of motion (2), the latter overlapping with *onne*:

- (1) *cise ohta an monimahpo nah kii manu*
house LOC exist.SG girl QUOT do REP

‘The girl who was in the house said this.’ — Murasaki 2001: text 52; West Sakhalin, Odasu (Asai Take)

- (2) *nay ohta san-i ike taa wahka kuu*
river LOC go.down.SG-NMLZ and EMPH water drink

‘He went down to the river and drank water.’ — Murasaki 2001: text 34; West Sakhalin, Odasu (Asai Take)

Temporal noun phrases take the same particle: *sihtoono ohta* ‘at midday’ (Murasaki 2001: text 46), *sine cup okta* ‘in one month’, *nuca kotan koro okta* ‘when the Russians held the country’ (Tangiku & Ogihara 2001: Letter 1).

The bare particle *ta* is no longer free with ordinary nouns: in the corpus it appears fused with place roots and relational nouns — *teeta* ‘here’, *taata* ‘there’, *kotanta* ‘in the village’, *ciseta* ‘at the house’ (17 tokens), and the *kasketa*-type forms of §11.8 — while full noun phrases are hosted by *ohta*. This distribution, which we draw from the corpus, matches Murasaki’s teaching rule: place names and common nouns take *ohta*, locative nouns take *-ta* (Murasaki 2025: 17); Chiri’s entry for *ta* likewise has it on roots and formal nouns (Chiri 1942: §123(150)).

A Sakhalin innovation extends *ohta* beyond location: Chiri documents, chiefly from Taraytomari, an overt object-marking use with effected objects and action nouns, where Hokkaidō would use plain word order Chiri (1942: §68):

- (3) *matutah homa usey ohta karasi manu*
woman.PL roe broth LOC make.PL REP

‘The women prepared the roe broth.’ — Chiri 1942: §68; West Sakhalin, Taraytomari

Likewise *ukoyki ohta kii* ‘they had a fight’, *hatto ohta kii* ‘he issued a ban’ (same section). Chiri suggests the construction grew out of the locative — ‘did (work) on the broth’.

11.3 Allative *onne* and *ene*

onne marks direction and endpoint, ‘to, toward’, and with verbs of entering ‘into’ Chiri (1942: §123(129)), (Murasaki 2025: 21):

- (4) *taa monimahpo wahka onne ahun manu*
EMPH girl water to enter.SG REP

‘The girl fell into the water.’ — Murasaki 2001: text 25; West Sakhalin, Odasu (Asai Take)

- (5) *ehopenpa kusu wahka onne repun manu*
be.startled because water to go.seaward REP

‘Startled, it plunged off into the water.’ — Kitahara et al. 2003; Sakhalin (dialect not further specified), Yamada Hayo

With verbs of perception the *onne*-phrase names the place a stimulus comes from — directionality construed from the perceiver:

- (6) *ne’an cise onne taa hemata yuukara hawe-he an manu*
that house to EMPH what song voice-POSS exist.SG REP

‘From inside the house there came the voice of some song.’ — Murasaki 2001: text 19; West Sakhalin, Odasu (Asai Take)

We analyse *onne* here as orientational ‘in the direction of’; the same frame recurs with *humih* ‘sound’ and *ene*, cf. (18).

The shorter allative *ene* — Chiri’s Sakhalin entry ‘...へ’ (Chiri 1942: §123(103)) — attaches in the corpus almost always to relational nouns (*empoke-ene* ‘to under’, *kaske-ene* ‘onto’, *onnayke ene* ‘to inside’) and to deictics (*hekimoh ene* ‘toward the hills’), not to plain nouns; examples follow in §11.8. The interrogatives pattern with the particles: *nakene* ‘whither’, *nahta* ‘where’, *nahwa* ‘whence’ (Murasaki 2025: 21).

11.4 Ablative *wa*, *orowa* and terminative *pahno*

The ablative proper is *wa* ‘from’, segmentally identical to the converb particle ‘and’ (Chapter 23) but attached to noun phrases Chiri (1942: §123(157)). After a nasal it assimilates: *kim-ma* ‘from the mountains’ (*kim wa*), a fusion frequent enough (53 corpus tokens) that we treat it as lexicalised:

- (7) *turano* *taa* *unkayoh* *kim-ma* *san* *manu*
 at.the.same.time EMPH ogre mountain-from come.down.SG REP

‘At that same moment the ogre came down from the mountains.’ — Murasaki 2001: text 2; West Sakhalin, Odasu (Asai Take)

Clause-initial *turano* is the independent use of the comitative, §11.6.

- (8) *muhru* *empoke* *wa* *tani* *neya* *kosonto* *soma-ha* *taa* *asinke* *manu*
 pillow underside from now that silk.robe bundle-POSS EMPH take.out REP

‘Then he took the bundle in the silk robe out from under the pillow.’ — Murasaki 2001: text 52; West Sakhalin, Odasu (Asai Take)

The expanded form *orowa* (‘from the place of’, *oro-wa*) does the same work with concrete sources and is the regular marker of the agent in passive-like indefinite constructions (Chapter 13) Chiri (1942: §123(137)); in Sakhalin the bare relational noun *oro* can stand for *oro wa* (Chiri 1942: §79).

- (9) *monimahpo* *taa* *uhsoro* *orowa* *taa* *sehpa* *sineh* *asinke* *ike* *horokewpo* *koore*
 girl EMPH bosom from EMPH sword.guard one take.out and man give

‘The girl took a sword-guard out of her bosom and gave it to the man.’ — Murasaki 2001: text 25; West Sakhalin, Odasu (Asai Take)

The terminative *pahno* ‘until, as far as’ bounds a path, a time span, or a quantity: *kunne pahno* ‘until dark’ (Murasaki 2001: text 19), *tani pahno* ‘until now’ (Murasaki 2001: text 4), quantitative ‘as much as’ in example (11) of Chapter 10. The East-coast letters write it *pak no*:

- (10) *israh* *kara* *wa* *cokay* *e=an* *toko* *pak* *no* *paye=an* *kusu* *iki*
 message make and 1PL 2SG=exist.SG place until ADV go.PL=INDEF.S PURP do

‘I will send word, and we will travel as far as the place where you are.’ — Tangiku & Ogihara 2001: Letter 1; East Sakhalin, Otasan (Sentoku Tarōji)

Path through or along something is marked by the prosecutives *kaari* ‘through, via’ and *okakara* ‘along’ (Chiri 1942: §123(109), (127)): *apa uturu kaari* ‘through the gap of the door’ (Murasaki 2001: text 4), *kina tun peka* ‘through the grass’ with the areal *peka* (Kitahara et al. 2003), and:

- (11) *kotan-* *ruwe-* *ohta* *oman* *teh* *tani* *taa* *ruu* *okakara* *oman* *ike* *taa* *ciseta* *oman*
 village- road- LOC go.SG and now EMPH road along go.SG and EMPH house- go.SG
 POSS POSS LOC

‘He took the road to his village, went along the road, and came to the house.’ — Murasaki 2001: text 25; West Sakhalin, Odasu (Asai Take)

Three markers in one route description: *ohta* for the goal of the first leg, *okakara* for the path, fused *-ta* for the final goal.

11.5 Instrumental *ani*

ani marks instruments and materials and, by extension, cause; it is the Sakhalin (and Hidaka) counterpart of *ari* elsewhere Chiri (1942: §123(98)). Murasaki glosses it ‘by means of’ and her dialogues use the formula *aynu itah ani* ‘in Ainu’ Murasaki (2025: 13) — the same phrase with which Asai Take’s interviewers begged her back from Japanese (*aynu itah ani yee kanne* ‘say it in Ainu’, (Murasaki 2001: text 37)).

- (12) *sine* *nii* *tah* *mukara* *ani* *humpa* *turano* *tah* *hemata* *hawe-he* *an* *manu*
 one tree that axe INS chop as.soon.as that what voice-POSS exist.SG REP

‘The moment she chopped at a tree with the axe, there was some kind of voice.’ — Murasaki 2001: text 49; West Sakhalin, Odasu (Asai Take)

Here *ani* is the instrumental on the noun *mukara* ‘axe’; the immediate-sequence marker is *turano* after the verb *humpa* ‘chop’ — for *turano* on a verb, see §11.9 on category overlap.

Causal *ani* — ‘because of this’ (*tampe ani*) — is recorded by Chiri from Pilsudski’s texts (Chiri 1942: §123(98)). In the Sentoku letters the instrument reading covers writing materials: *pon cicay ani kanpi* ‘a letter (written) on a small scrap’ (Tangiku & Ogihara 2001: Letter 1).

11.6 Comitative *tura* and *turano*

tura ‘together with’ is at bottom a transitive verb ‘accompany’, and much of its corpus life is verbal (*tura ahun* ‘went in with (her)’, *tura makan* ‘went up together’); postposed to a noun phrase it is the comitative case marker, optionally extended with adverbial *-no*, *turano* Chiri (1942: §123(153–154)).

- (13) *re horokewpo tura an manu*
three man COMIT exist.SG REP

‘She lived together with the three young men.’ — Murasaki 2001: text 51; West Sakhalin, Odasu (Asai Take)

Like the other particles, it has an anaphoric use without a host — example (7) above, *turano* ‘simultaneously with that’ — and a use with verbs, ‘while V-ing, as soon as V’ (*cis turano* ‘while crying’, *humpa turano* in (12)); Chiri registers both extensions (Chiri 1942: §123(153)). In the letters, *tura* coordinates noun phrases outright: *Seraroka aynu Sakayhama aynu tura* ‘the Seraroka people together with the Sakayhama people’ (Tangiku & Ogihara 2001: Letter 1). The reciprocal-marked adverb *uturano* ‘all together’ belongs to the same family (Murasaki 2025: 21).

11.7 Mutative *ne*

The copula *ne* ‘be’ (Chapter 18), used without person marking after a noun phrase, functions as a case particle of state: ‘as X’, ‘(turn) into X’ Chiri (1942: §123(121)). The commonest frames are with *kara* ‘make (into)’, *koro* ‘take as’ (*mah ne koro* ‘take as wife’, example (4) of Chapter 7), and verbs of becoming:

- (14) *seta haw ne kara teh taa hohpa teh oman manu*
dog voice COP make and EMPH leave and go.SG REP

‘He made his voice into a dog’s voice and went off, leaving her.’ — Murasaki 2001: text 46; West Sakhalin, Odasu (Asai Take)

Chiri further records for Sakhalin a directional use of *ne* equivalent to *ta*, as in *soy ne e-yah* ‘you threw it outside’, and a temporal one (*numan-ne* ‘yesterday’) Chiri (1942: §123(121)); in the corpus this survives in fossilised adverbs (*hekimoh ne ~ hekimoh ene* ‘toward the hills’), and we take Chiri’s *ne* here to be the reduced form of allative *ene*.

11.8 Relational nouns

Fine-grained spatial relations are encoded by a closed class of bound nouns that follow a reference-point noun phrase: the construction is [N (+POSS) + relational noun + case]. Chiri defines the class by its person marking: a relational noun cannot take the possessive suffix of Chapter 7 in its spatial use, and a pronominal reference point appears as an *object* prefix — *en-ka* ‘above me’, *i-ka* ‘above us’, never **an-kasi* ‘my top’ Chiri (1942: §§78–79). Murasaki teaches the same pattern for Rayciska with the accusative person prefixes: *en=empokehe* ‘below me’, *en=oponi* ‘after me’ Murasaki (2025: 26); and the reciprocal prefix yields *u-sam* ‘next to each other’ (Dal Corso 2024: 74) — built on the relational noun *sam* ‘side’, not the homophonous verb *sam* ‘marry’ whose reciprocal *u-sam* means ‘marry each other’ (Chapter 14).

The inventory for Rayciska, with the locative *-ta* attached, is listed by Murasaki Murasaki (2025: 17–18): *kaske* ‘top (in contact)’ and elders’ *kaa*; *enkaske* ‘above (detached)’ and elders’ *enkaa*; *kitayke* ‘summit, source’; *empoke* ‘underside’; *sanke* ‘vicinity’; *tunke* ‘inside (a mass)’; *noske* ‘middle’; *onnayke* ‘inside (an enclosure)’; *soyke* ‘outside’. Chiri’s island-wide list adds among others *corpoke* ‘under’, *etoko* ‘tip, front, before’, *sam* ‘beside’, *oka(ke)* ‘behind, after’, *uturu* ‘interval’, *caa* ‘edge, bank’ (Chiri 1942: §79). Most are built on monosyllabic roots (*ka* ‘up’, *pok* ‘down’, *sam* ‘side’) extended by *-ke*, which Chiri analyses as rhythmic rather than meaningful (Chiri 1942: §79).

With the locative the relational noun takes *-ta* directly:

- (15) *yan nii kasketa ama manu*
come.ashore tree top-LOC put REP

‘He put it on top of a driftwood log.’ — Murasaki 2001: text 5; West Sakhalin, Odasu (Asai Take)

The same passage has the elders’ variant *yan nii kaata* ‘on the driftwood’, exactly the *kaske ~ kaa* doublet Murasaki reports.

- (16) *cise onnayke-ta sine monimahpo an teh ...*
house inside-LOC one girl exist.SG and ...

‘Inside the house there lived a girl, and...’ — Murasaki 2001: text 41; West Sakhalin, Odasu (Asai Take)

Direction takes *ene*, and source takes *wa*, as already seen in (8); the three cases line up on one and the same relational noun, e.g. *empoke-ta* ‘under’ (17), *empoke ene* ‘to under (the pillow)’ (Murasaki 2001: text 52), *empoke wa* ‘from under’ (8):

- (17) *e=hoku-hu* *sookuma* *empoke-ta* *an* *teh* *kaske-ene* *e=okoyse* ...
 2SG=husband-POSS rock.island underside-LOC exist.SG and top-to 2SG=urinate ...

‘Your husband is under the rock island, and you urinate down on top of him...’ — Murasaki 2001: text 28; West Sakhalin, Odasu (Asai Take)

kaske-ene with no overt host noun: the reference point (‘him’) is anaphoric, third person being zero-marked.

- (18) *cise* *onnayke* *ene* *hemata* *humi-hi* *an* *manu*
 house inside to what sound-POSS exist.SG REP

‘From inside the house there was some kind of sound.’ — Murasaki 2001: text 4; West Sakhalin, Odasu (Asai Take)

The perceiver-oriented frame of (6) again, here with a relational noun.

Temporal uses follow the spatial ones: *etoko ta* ‘before’ (*cep koyki etoko ta* ‘before the fishing season’, (Tangiku & Ogihara 2001: Letter 1); *iine too etoko ta* ‘four days before’, (Chiri 1942: §79)) and *okaaketa* ‘after that’, ubiquitous in Asai Take’s narration (Murasaki 2001: text 52).

When a relational noun is promoted to a referring noun — ‘the top’, ‘the inside’, ‘the tip’ as things — it rejoins the ordinary noun class and takes possessive morphology after all: Chiri notes substantivised *kasike* ‘face’, *oske* ‘heart’ (Chiri 1942: §78), Murasaki gives *cih onnaykehe poro* ‘the inside of the boat is big’ (Murasaki 2001: text 17), and the corpus has:

- (19) *makiri* *asinke* *ike* *etoko-ho* *kayteh* *teh* *an* *manu*
 knife take.out and tip-POSS break and exist.SG REP

‘He took out his knife, and its tip was broken off.’ — Murasaki 2001: text 52; West Sakhalin, Odasu (Asai Take)

11.9 Particle, relational noun, or clause linker?

The three categories of this chapter shade into one another, and several high-frequency items live on the boundary. The diagnostics we apply are host class (noun phrase vs verb/clause) and person marking (object prefixes mark relational nouns; particles take none).

kusu ‘because; in order to’ is in origin one of Chiri’s ‘second-class formal nouns’ (*kus* ~ *kusu*) (Chiri 1942: §78), but synchronically it attaches to clauses, marking cause or purpose (Chiri 1942: §121(67)) — as in (5), (10) — and to interrogatives (*hemata kusu* ‘why’, (Murasaki 2001: text 19)); clause linking is treated in Chapter 23.

renkayne is listed by Chiri twice, as a clause-linking particle ‘because, thanks to’ and as a case particle on nouns Chiri (1942: §121(79), §123(145)). The East-coast letters use the clausal value:

- (20) *tani* *nuca* *isan* *renkayne*
 now Russian not.exist because

‘because the Russians are gone now’ — Tangiku & Ogihara 2001: Letter 1; East Sakhalin, Otasan (Sentoku Tarōji)

In Asai Take’s West-coast tales, by contrast, *renkayne* is a quantifier ‘in numbers, plenty’ (Chapter 10), and the causal value is carried by *kusu* and *ani*. We analyse the two as one lexeme (ultimately *renkay* ‘will, consent’ plus *ne*, cf. *an-renkayne* ‘according to our wish’ in Dobrotvorsky’s texts, (Dobrotvorsky 1875); (Sakaguchi 2021)) that grammaticalised along two routes — ‘by the will of, thanks to’ in the east, ‘to one’s heart’s content’ > ‘in plenty’ in the west. A dialect-split polysemy of this kind would be settled by causal tokens in West-coast narrative; we have found none in the corpus.

Finally, the particles themselves recruit verbal hosts. Instrumental *ani* on an infinitive verb yields ‘by V-ing’ (Chiri’s Sakhalin example *kira ani eh* ‘came fleeing’, (Chiri 1942: §123(98))), and comitative *turano* on a verb means ‘as soon as’, (12). The converse also holds: ablative *wa* and the converb *wa* ‘and’ (Chiri 1942: §121(88)) are segmentally one item distinguished only by host class. Sakhalin Ainu thus presents a single postposed relator slot whose exponents are distributed over nominal and verbal hosts, with category labels — case particle, relational noun, conjunctive particle — naming positions on a grammaticalisation path rather than watertight classes.

Part IV · The verb

Chapter 12 · Verb structure and transitivity

The verbal template, strict transitivity classes, verbal number, and suppletion.

The verb is the largest and most elaborated word class of Sakhalin Ainu (Chapter 6) and the only obligatory constituent of the clause: *oman* ‘he went’, *pirika* ‘it is good’ are complete sentences (Chiri 1942: §21–22). A finite verb indexes its core arguments (Chapter 13), registers their number partly in its own stem and partly by a suffix, and changes its valency only by overt derivation (Chapter 14); it does not inflect for tense, temporal and modal meaning being carried by following auxiliaries and particles (Chiri 1942: §26) (Chapter 15). This chapter describes the shape of the verbal word and the lexical classification that the rest of the verbal grammar presupposes: the template, the transitivity classes, verbal number, stem classes, and the light verb *ki* ‘do’.

The descriptive foundation is Chiri (1942), whose Sakhalin-centred grammar classifies verbs by the arguments they require (his *complete* and *incomplete* predicates, §§27–33) and by their number behaviour (§§34–36); Murasaki (1979) covers the west-coast Rayciska variety, Tangiku (1998) the morphology of stems in *-rV*, and Dal Corso (2024) the east-coast verb as recorded by Pilsudski. A note on citation: verbs are cited in this grammar in their bare third-person form, which is also the stem — *oman* ‘go’, *nukara* ‘see’ — since the language has no infinitive; the bare form serves both as the third person and as the pure lexical concept (Chiri 1942: §22).

12.1 The verbal word and its template

The maximal verbal word of Sakhalin Ainu has the structure in (i): person prefixes or proclitics; one or more valency-changing prefixes (reciprocal *u-*, reflexive *yay-/si-*, antipassive *i-*, applicative *e-/ko-/o-*); an incorporated noun; the root; valency-changing suffixes (causative *-te/-re*, *-ke*, *-ka*); the desiderative *rusuy*; the person suffixes *-an*, *-as*, *-yan*; and, outside them, the plural clitic *=hci* (Chapter 8). Chiri demonstrated the prefixal expansion with built-up forms like *u-ko-itak* ‘talk to each other’ and *usa-oruspe a-e-u-ko-itak* ‘we talk to each other about various news’, where the applicative *ko-* licenses the reciprocal and *e-* introduces the topic of conversation (Chiri 1942: §25); the polysynthesis that made Ainu famous in typology is this machine at full stretch, though he is careful to note that such piling-up is rare outside epic diction (Chiri 1942: §25).

The suffixal field is rigidly ordered. Causative suffixes attach directly to the stem, before the echo vowel of *-rV* stems is supplied (*kor-te* → *konte* ‘give’; Chapter 5); the desiderative *rusuy* likewise reaches the bare stem (*an-kon rusuy* ‘I want to have’), and the plural suffix *-(a)hci* attaches outside it. Tangiku concludes from this ordering that *rusuy* is a suffix rather than a free auxiliary in Sakhalin Ainu (Tangiku 1998: 231–232):

- (1) *tura esine kotan ohta kayki reekoh an-kon rusuy-ahci yahka*
with.it same village in even very INDEF.A-have DES-PL although

‘Even from that same village many wanted to have her in marriage, but...’ — Murasaki 1976; Tangiku 1998: 227; West Sakhalin, Rayciska

The stem *kor-* loses its echo vowel before *rusuy* (*kon rusuy*), and the plural *-ahci* follows *rusuy*: stem-causative-rusuy-plural is the fixed order of the suffixal field.

Beyond the word so defined comes a short chain of postverbal auxiliaries and particles, unmarked for person, which carry aspect, modality and evidentiality. Table 1 gives the commonest ones with corpus attestations; their semantics is the business of Chapter 15 and Chapter 16.

form	meaning	corpus example	source
<i>rusuy</i>	desiderative	<i>an-kon rusuy-ahci</i> ‘they wanted to have’	(Tangiku 1998: 227)
<i>rayki</i>	‘long to’	<i>mokon rayki</i> ‘be dying to sleep’	(Hattori 1964); (Tangiku 1998: 228)
<i>kusu kara</i>	‘be about to’	<i>mokoro kusu kara</i> ‘was about to sleep’	(Murasaki 2001: text 46)
<i>kusu an</i>	progressive	<i>hecirehci kusu an</i> ‘were dancing’	(Murasaki 2001: text 46); (Dal Corso 2024: 72)
<i>kusu iki</i>	‘be going to’	<i>rayki kusu iki</i> ‘will catch’	(Tangiku & Ogihara 2001: Letter 1)
<i>hemaka</i>	completive ‘finish’	<i>ray hemaka</i> ‘died (for good)’	(Murasaki 2001: text 27)
<i>manu(u)</i>	reportative	<i>san manu</i> ‘went down, they say’	(Murasaki 2001)
<i>nanko(ro)</i>	‘probably; surely’	<i>an-ki nanko</i> ‘we shall do it’	(Dobrotvorsky 1875); (Sakaguchi 2021)
<i>yan</i>	plural imperative	<i>toh roski yan</i> ‘set up the bamboo!’	(Murasaki 1979: 30); (Murasaki 2001: text 8)

Table 25: Postverbal auxiliaries and particles (overview)

The auxiliaries are person-blind: in the corpus only the plural *-(a)hci* ever intervenes onto them (*rusuy-ahci* above), and person markers stay on the lexical verb or on a supporting *ki* ‘do’ (§12.5). Where the literature is silent on *kusu kara* and *hemaka*, the statement above is our generalisation from Asai Take’s tales, where both are frequent and never person-marked.

12.2 Transitivity

Every verb stem of Sakhalin Ainu is lexically specified for its valency. Chiri’s classification divides predicates into *complete* (requiring no complement: intransitives, including all property verbs) and *incomplete* (requiring one: transitives, plus the copula *ne*, the single underived verb that takes a complement rather than an object) (Chiri 1942: §§27, 32). The diagnostic is morphological, and it is the pivot of the person system: a first-person or indefinite subject is suffixed to a complete verb but prefixed to an incomplete one — *ahun-an* ‘I went in’ against *an-ahunke* ‘I put it in’ (Chiri 1942: §28) (Chapter 13). Homophonous stems are kept apart by exactly this test (Chiri 1942: §29):

stem	intransitive	transitive
<i>ma</i>	<i>ma-an</i> ‘I swam’	<i>an-ma</i> ‘I roasted it’
<i>kay</i>	<i>kay-an</i> ‘I broke (in two)’	<i>an-kay</i> ‘I carried it on my back’

Table 26: Homophonous intransitive–transitive pairs (Chiri 1942: §29)

There is no labiality. A stem never alternates between transitive and intransitive use without overt marking; valency is adjusted only by the derivational morphology of Chapter 14 — compare *asin* ‘go out’ with *asinke* ‘take out’, *ahun* ‘enter’ with *ahunke* ‘put in’ (Chiri 1942: §§33, 55). Where English uses one verb in ‘the door opened’ and ‘I opened the door’, Sakhalin Ainu pairs *makke* with *maka*, both derived from the bare root *mak-* (Chiri 1942: §55) (§12.4). Even lexical semantics that straddles the divide is split over two stems: ‘think’ is complete *yaynu* with quoted thought but incomplete *ramu* with a complement clause, as ‘say’ is *itah* against *ye* and ‘eat’ is *ipe* against *ee* ‘eat’ (Chiri 1942: §30). Chiri knew only two verbs used in both classes, *roski* ‘stand (pl)’ and *kari* ‘turn’ (Chiri 1942: §31).

12.2.1 Ditransitives Three-place predicates are derived, not basic: the recipient verbs are causatives of possession and ingestion. *konte* ‘give’ is *kor-te* ‘cause to have’, and Chiri insists that in *tampe huci ku-konte* ‘I give this to grandmother’ (his §68 example, cited in pan-Ainu shape as *ku-kore*) the grandmother is the causee-object, not a dative (Chiri 1942: §68). The recipient is the argument the verb indexes; the theme stays a bare unindexed noun:

(2) *сiне котан короуа иконде*

sine kotan koro wa i=konte
one place have and INDEF.O=give

‘Grant me one fishing ground to hold.’ — Tangiku & Ogihara 2001: Letter 1; East Sakhalin, Otasan (Sentoku Tarōji)
The recipient (‘me’) is indexed by *i=*; *kotan* ‘village, place’ here denotes a fishing station.

(3) *suke-hci ike taa tani tah horokeypo ee-re tura e-hci*
cook-PL and EMPH now this man eat-CAUS together.with eat-PL

‘They cooked them, served them to the man, and ate together with him.’ — Murasaki 2001: text 5; West Sakhalin, Odasu (Asai Take)
eere ‘feed’, from *ee* ‘eat’ with the causative *-re*, takes the fed person as its indexed object.

12.2.2 Weather and ambient verbs At the bottom of the valency scale stand zero-valent predicates of weather and ambience. The productive type incorporates *sir* ‘ambient scene’ — one of Chiri’s desemanticised ‘formal nouns’ (Chiri 1942: §78) — into a property verb: *siripirika* ‘the weather is fine’, *siriwen* ‘the weather is bad’, with the regular Sakhalin echo vowel on *siri-*. In his survey of Ainu noun incorporation Satō counts such natural-force subject incorporation as the second commonest type after object incorporation (Satō 2022: ch. 16, in (Bugueva 2022)), and Dal Corso identifies east-coast *sistono* ‘day breaks’ as a zero-valency verb of natural conditions (Dal Corso 2024: 73).

(4) *tanto ampene siri-pirika*
today completely weather-be.good

‘Today the weather is thoroughly fine.’ — Murasaki 2025; West Sakhalin

Beside the incorporated type stands an analytic pattern with a weather noun as ordinary subject of a motion or property verb: *ahto ran* ‘rain falls’, *reera yuhke* ‘the wind is strong’ (Chapter 6), *onuman oman* ‘evening comes’. The two strategies coexist in the same texts; we find no meaning difference beyond the referentiality of the weather noun, which in the analytic type can be modified and quantified (*poro ahto ran* ‘a great rain fell’).

(5) *onuman oman teh tani taa reekoh ahto ran manu*
evening go.SG and now EMPH very rain descend.SG REP

‘Evening came on, and it rained very hard.’ — Murasaki 2001: text 24; West Sakhalin, Odasu (Asai Take)

12.3 Verbal number

A closed set of verbs distinguishes singular and plural stems. Number here is the number of the event’s absolutive participant — the subject of an intransitive, the object of a transitive — and, more exactly, of the event itself: one entering against many enterings (Chiri 1942: §§34–36), (Sakaguchi 2024: 124). Three formal types occur in Sakhalin (Chiri 1942: §35), (Sakaguchi 2024: 122–123): suppletion (Table 3); the *-n/-p* alternation of seven motion verbs (Table 4); and a residual class with plural *-pa* (*tuye ~ tuypa* ‘cut’, *nasa ~ naspa* ‘split’, (Murasaki 1979: 42)), which is far less productive than in Hokkaidō — most Hokkaidō *-pa* plurals have no Sakhalin counterpart, the plural suffix *-hci* doing their work (Tangiku 2022), (Sakaguchi 2024: 123).

SG	PL	meaning	attestation
<i>an</i>	<i>okay</i>	‘exist, be at’	(Chiri 1942: §35); (Sakaguchi 2021); (Murasaki 2001: text 4)
<i>oman</i>	<i>paye(h)</i>	‘go’	(Chiri 1942: §35); (Murasaki 2001: text 51)
<i>eh (ek-)</i>	<i>ariki(h)</i>	‘come’	(Sakaguchi 2021); (Tangiku & Ogihara 2001: Letter 3); (Sakaguchi 2024: 122)
<i>as, hotari</i>	<i>roski</i>	‘stand (up)’	(Chiri 1942: §§31, 35); (Sakaguchi 2024: 122); (Murasaki 2001: text 8)
<i>numa</i>	<i>payke</i>	‘get up’	(Sakaguchi 2024: 122)
<i>uh (uk-)</i>	<i>uyna</i>	‘take’ (object number)	(Chiri 1942: §36); (Kitahara et al. 2003); (Murasaki 2001: text 51)

Table 27: Suppletive stem pairs attested in Sakhalin Ainu

Hokkaidō's pair *rayke* ~ *ronnu* 'kill' does not belong here: in Sakhalin *rayki* ~ *rayke* are regional variants, *ronnu* is recorded only as a stylistic alternative, and the Sakhalin prose corpus has no number opposition between them (Hattori 1964: 29), (Sakaguchi 2024: 123). In the corpus we likewise find no token of *ronnu*.

SG	PL stem	meaning	corpus
<i>ahun</i>	<i>ahup-</i>	'go in'	<i>ahup-ahci</i> (Murasaki 2001: text 5)
<i>asin</i>	<i>asip-</i>	'go out'	(Chiri 1942: §35)
<i>san</i>	<i>sap-</i>	'go down'	<i>sap-ahci</i> (Murasaki 2001: text 19)
<i>makan</i>	<i>makap-</i>	'go up(hill)'	<i>makap-ahci</i> (Murasaki 2001: text 37)
<i>rikin</i>	<i>rikip-</i>	'climb'	<i>rikip-ahci</i> (Murasaki 2001: text 52)
<i>yan</i>	<i>yap-</i>	'go ashore'	<i>yah, yap-an</i> (Tangiku 1998: 223)
<i>ran</i>	<i>rap-</i>	'come down'	(Chiri 1942: §35)

Table 28: The n/p motion verbs (Chiri 1942: §35; corpus forms)

The bare plural stems of Table 4 end in a stop and therefore surface with the coda *h* when nothing follows (*yah* 'they went ashore'); before vowel-initial suffixes the stop reappears (*yap-an*, *makap-ahci*; Chapter 5). In the corpus the bare plurals are rare: they are almost always clothed in *-ahci* or a person suffix, an observation we owe to our own counts.

- (6) *Охзороно окай анъ тренъкайнэ, тәне танъ котанъ та эх.*

ohorono okay=an renkayne tane tan kotan ta eh
 long.time exist.PL=INDEF.S because now this village in come.SG

'Since I stayed on a long while, now you have come to this village.' — Dobrotvorsky 1875; Sakaguchi 2021; West Sakhalin, West coast, central (recorded 1867–1872)

The indefinite person, grammatically plural, selects the plural stem *okay* for a single speaker; the third-person singular addressee takes singular *eh*.

Stem number and the plural suffix *-hci* are independent systems that stack. *-hci* marks plurality or collectivity of a core participant on any verb (Chapter 8; for its interaction with person, Chapter 13); on a verb with stem number it attaches to the appropriate stem, so plural reference is regularly marked twice:

- (7) *monimahpo taa reh pis taa utura-hci ahup-ahci manu*
 girl EMPH three piece EMPH be.together-PL enter.PL-PL REP

'The three girls went in together.' — Murasaki 2001: text 5; West Sakhalin, Odasu (Asai Take)

With transitive members of the set, stem number tracks the object while *-hci* or *yan* tracks the subject independently: *uyna-hci* 'they gathered them up' (Murasaki 2001: text 51), and in the plural imperative:

- (8) *toh roske-yan toh roske-yan*
 bamboo stand.PL-IMP.PL bamboo stand.PL-IMP.PL

'Set up the bamboo, set up the bamboo!' — Murasaki 2001: text 8; West Sakhalin, Odasu (Asai Take)

roske, the plural of *asi* 'stand one thing up', agrees with the many bamboos; *-yan* marks the plural addressees.

Sakaguchi notes a single east-coast singular 'come', *eki* ~ *iki*, used in Yamabe's narrative for plural subjects as well, and even combined with the plural suffix (*eki-hsi*) (Sakaguchi 2024: 122) — a hint that for 'come' the suppletion was already decaying in the early twentieth century, while the west-coast tales of Asai Take keep all the pairs of Tables 3–4 intact.

12.4 Stem classes

For the morphology of the preceding sections, four stem classes need to be distinguished. (i) Vowel-final stems: *ee* 'eat', *tura*, *kira* 'flee', *tuuri* 'stretch'. They take post-vocalic allomorphs throughout: causative *-re*, plural *-hci*. (ii) Consonant-final stems whose final stop or *m* is neutralised word-finally: *itah* 'speak' (*itak-*), *ahkas* 'walk', *ahun*. They take *-te*, *-ahci*, and their underlying coda reappears before vowels (*itak-ahci*, *yap-an*; Chapter 5). (iii) Stems in *-rV*, which split in two: underlying *r*-final stems (*koro* 'have', *nukara* 'see', *mokoro* 'sleep') behave as consonant-final under suffixation — *konte*, *nukante*, *kon rusuy*, *nukan rusuy* — while true vowel-final ones

(*poro*, *kira*, *tuuri*) do not; this is Tangiku's demonstration, set out with the phonology in Chapter 5 (Tangiku 1998: 225–230). The same *r* may neutralise to coda *h* instead of taking the echo vowel — *utah* beside *utara* ‘people’, *mahnkeh* ‘woman’ (Chiri 1942: §16), (Dal Corso 2024: 64). (iv) Bare roots of shape CVC that surface only with a derivational suffix: *mak-* ‘open’ in *mak-a* (transitive, thematic vowel) and *mak-ke* (intransitive), *per-e* ‘tear’ beside *per-ke* ‘be torn’ (Chiri 1942: §55) — the seedbed of the valency pairs of Chapter 14.

- (9) *otaka ene san inkan rusuy omantene tani*
beach to go.down.SG ANTIP.see DES and.then now

‘He went down to the beach and, wanting to take a look, then...’ — Tangiku 1998: 227; West Sakhalin, Odasu (Asai Take)

inkara ‘look’ (from *i-nukar*) is an underlying *r*-stem: *inkan rusuy*, like *kon rusuy*.

Citation forms in this grammar follow the surface word: *-rV* verbs are cited with their echo vowel (*nukara*, not *nukar*), coda-neutralised verbs with *h* (*itah*, *yah*), and suppletive verbs as a pair with the singular first (*oman* ~ *paye*). Where an underlying stem matters we add it in parentheses (*itak-*, *kor-*).

12.5 The light verb *ki*

The verb *ki* ‘do’ — Sakhalin *kii* with the regular lengthening of monosyllables, shortened again under suffixation (*ki-hci*) (Tangiku 2022) — is the language's all-purpose support verb. Chiri's classic diagnostic for the unity of verbs and property words turns on it: focused predicates of either kind are resumed by *ki*, as in *oman ka ki* ‘he did go’ and *an ka somo ki* ‘he is not there at all’ (Chiri 1942: §21) (Chapter 6); negation with a focus particle takes the same shape, *hosipi kayki han ki* ‘he did not so much as return’ (Pilsudski 1912: 13), (Chiri 1942: §49) (Chapter 17).

Beyond focus support, *ki* is a dummy predicate over a lexical verb or verbal noun, which keeps its own internal morphology while *ki* carries the clausal trappings. Dobrotvorskij's nineteenth-century conversations build whole exchanges on it:

- (10) *оровано укораму пирика анъ ки нанго*.

orowano u-ko-ramu-pirika an-ki nanko
from.then RECP-APPL-mind-be.good INDEF.A-do probably

‘From now on we shall make peace with each other.’ — Dobrotvorskij 1875; Sakaguchi 2021; West Sakhalin, West coast, central (recorded 1867–1872)

The complex predicate *ukoramupirika* stays bare; person and modality ride on *an-ki nanko*.

- (11) *нуця исамике текоро рамисахсе анкихи*

nuca isam-ike tekoro ramisahse an=ki-hci
Russian not.exist-and very feel.lonely INDEF.A=do-PL

‘With the Russians gone, we are very lonely.’ — Tangiku & Ogiwara 2001: Letter 1; East Sakhalin, Otasan (Sentoku Tarōji)

The construction can even split the person marking of a single semantic predicate: in Yamabe's autobiography the object person sits on the lexical verb and the subject person on *kii*, the lexical verb serving as the dummy's object (Sakaguchi 2024: 52):

- (12) *anoka ne anahkayki eciokay eci-ueomante hane an-kii no*
INDEF COP although 2PL 2PL.O-worry.about NEG INDEF.A-do and

‘And I, for my part, (set out) without worrying about you either.’ — Sakaguchi 2024: 52, ex. 3-48; East Sakhalin, Yamabe Yasunosuke

Finally, *ki* forms lexicalised activity predicates with a bare noun: *ceh ki* ‘fish’ (literally ‘do fish’), *yuh ki* ‘hunt’ (‘do game’), beside the quotative pro-verb *nah ki* ‘do thus’. Asai Take strings them into the formula that opens many of her tales:

- (13) *ceh kii koh ee niina koh kuru ehah taa koh ee nah kii yayne*
fish do whenever eat gather.firewood whenever burn lily.root dig whenever eat QUOT do while

‘She would catch fish and eat, gather firewood and burn it, dig lily roots and eat — and as she lived on so...’ — Murasaki 2001: text 39; West Sakhalin, Odasu (Asai Take)

The same collocations occur unverbated, with the plural suffix treating noun plus verb as one word (*cehki-hci* ‘they fished’, (Murasaki 2001: text 4)) — the bridge into noun incorporation, taken up in Chapter 14. The variant *iki* appears in the future periphrasis *kusu iki* (Table 1) and as a free variant of *ki* in the east-coast letters ((Tangiku & Ogihara 2001: Letter 1)).

Chapter 13 · Person marking and alignment

The person-marking paradigms, the indefinite person set, portmanteau prefixes, and morphological alignment.

Sakhalin Ainu is a head-marking language: a finite verb indexes the person of its core arguments by affixes, and independent pronouns (Chapter 9) appear only for emphasis or contrast. The inventory of person markers is small — five overt series and a zero third person — but their combinations in transitive verbs are intricate, and they have been a centre of discussion since the earliest descriptions. The foundations are Chiri (1942: §37), Hattori (1961) and Murasaki (1979: 49–51) for the west-coast Rayciska variety, and Satō (1985) for the east coast; Sakaguchi (2024) gives the fullest philological account of the east-coast paradigm, and Dal Corso (2025a) the most recent analysis of its alignment and history. This chapter follows the same order: the basic paradigm, the indefinite set, object marking, the first–second person interaction forms, and the alignment they jointly define, with short sections on nouns and on verbal number.

A word on terminology. Beside first and second person, Sakhalin Ainu has the set *-an / an- / i-*, called the *fourth person* in Japanese-language scholarship, *indefinite person* since Tamura's work on Saru, *inclusive* by Bugaeva, and *participant referentiality mismatch marker* by Dal Corso (Dal Corso 2025a: 196–200). We gloss it INDEF throughout. Following Satō (1985) and Sakaguchi (2024: 26–28), it is also convenient to speak of three first-person *series* named after their transitive subject form: the KU-series (*ku-*, pronoun *kuani* ‘I’), the CI-series (*ci-/as*, pronoun *ciokay* ‘we’ ~ *cookay*), and the AN-series (the indefinite set, pronoun *anoka* ‘I, we’ ~ *anokay*).

13.1 The person-marking paradigm

Table 1 sets out the person markers of the Sakhalin verb. It combines Chiri (1942: §37), Murasaki (1979: 49–50), Satō (1985), Sakaguchi (2024: 29–34) and Dal Corso (2025a: 200–201); where the coasts diverge, the western form is marked W and the eastern E.

	S	A	O
1SG	<i>ku-</i>	<i>ku-</i>	<i>en-</i> (W) ~ <i>in-</i> (E)
1PLEXCL	<i>-as</i>	<i>ci-</i>	<i>un-</i> (W, rare) ~ <i>in-</i> (E)
2SG	<i>e-</i>	<i>e-</i>	<i>e-</i>
2PL	<i>eci-</i>	<i>eci-</i>	<i>eci-</i>
3	—	—	—
INDEF	<i>-an</i>	<i>an-</i>	<i>i-</i>

Table 29: Person markers on the Sakhalin Ainu verb

First person singular *ku-* prefixes to intransitive and transitive verbs alike. It is the everyday first person of conversation and of quoted dialogue in Asai Take's tales:

- (1) *hemata ku-kara-ha kuani neyahka ku-inkara kusu otakaa-ta ku-san*
 what 1SG.A-do-POSS 1SG also 1SG.S-look PURP beach-LOC 1SG.S-go.down.SG

‘I too came down to the beach to see what I should do.’ — Murasaki 2001: text 30; West Sakhalin, Odasu (Asai Take)

Second person singular *e-* and plural *eci-* are invariant for role; Chiri records *esi-* for 2PL at Tarantomari (Chiri 1942: §37). Sentoku Tarōji (千徳太郎治) writes *e=* to Pilsudski:

- (2) *e=koro kotan okta e=oman kusu nah e=ye renkayne*
 2SG.A=have village to 2SG.S=go.SG PURP QUOT 2SG.A=say because

‘because you said that you would go back to your own village’ — Tangiku & Ogihara 2001: Letter 1; East Sakhalin, Otasan (Sentoku Tarōji)

Third person is unmarked in all roles and numbers. With Chiri we take the absence of an affix on a finite verb to be itself interpretable — a zero exponent read as third person wherever no overt marker appears (Chiri 1942: §23, §37); plurality of a third person argument may be signalled by *-hci* (§13.7).

The CI-series splits by transitivity: *-as* suffixes to intransitives, *ci-* prefixes to transitives and the copula. It is productive on the east coast, while in Rayciska it is marginal — Hattori's consultants knew *cioka* as the speech of elders or of the southern west coast, and Murasaki's Rayciska corpus contains a single possessive token Hattori (1961: 5), (Sakaguchi 2024: 27). In the east-coast folklore the CI-series regularly refers to a single speaker: it is grammatically plural (it selects plural verb stems) but semantically ‘I’ (Satō 1985), (Sakaguchi 2024: 27). We keep the conventional gloss 1PL.EXCL while flagging this mismatch wherever it matters.

- (3) *nahte asip-as cikin paye-as sir-ihī eci-nukan-te*
 then go.out.PL-1PL.EXCL.S if go.PL-1PL.EXCL.S appearance-POSS 1PL.EXCL.A>2SG.O-see-CAUS

‘Then, when I go out, I will show you how I go.’ — Piłsudski 1912: text 13, ll. 24–25; Sakaguchi 2024: ex. 3–1; East Sakhalin

The she-bear speaks of herself alone; the CI-series forms nonetheless take the plural stems *asip* and *paye*, and the 1→2 interaction surfaces as the portmanteau *eci-* (§13.4).

- (4) *naa iani e=an manuy ci=nu renkayne*
 still 2SG 2SG.S=exist.SG REP 1PL.EXCL.A=hear because

‘because we heard that you were still there’ — Tangiku & Ogihara 2001: Letter 1; East Sakhalin, Otasan (Sentoku Tarōji)

The series do not mix freely. Satō (1985) formulated a *personal agreement rule*: within one clause (pronoun included), markers of the AN-series do not co-occur with markers of the KU- or CI-series, while KU- and CI-forms do co-occur with each other — Dobrotvorsky already records *cookay ku-eraman* ‘I understand’ with the CI-pronoun and the KU-prefix (Dobrotvorsky 1875: 150), (Sakaguchi 2024: 30). The rule is what justifies sorting the markers into series in the first place, and §13.4 puts it to work.

13.2 The indefinite person

The set *-an* (S), *an-* (A), *i-* (O), with the pronoun *anoka* ‘I, we’ ~ *anokay*, is the workhorse of Sakhalin first-person reference. Its attested functions are: a general first person plural ‘we’, with no exclusive–inclusive opposition of the Hokkaidō colloquial kind; an ordinary first person singular ‘I’; an indefinite or unspecified agent; and the agent slot of the impersonal passive construction (Chiri (1942: §37); Sakaguchi (2024: 27–28); Dal Corso (2025a: 196–197); Nakagawa (2024: 179), who likewise gives Sakhalin *an=* the first-person-plural slot and notes the near-absence of the Hokkaidō exclusive *ci=*). The honorific second-person use of Hokkaidō dialects is unattested in Sakhalin (Dal Corso 2025a: 196). Chiri observed that Sakhalin uses one and the same paradigm in literary and colloquial registers, essentially the epic paradigm of Hokkaidō (Chiri 1942: §37).

The indefinite is grammatically plural regardless of its reference: where a verb has suppletive number stems, the AN-series selects the plural stem even for a single referent (§13.7). In Sentoku's letters it is the unmarked way of saying both ‘I’ and ‘we’:

- (5) *anokay aynu utara ramma ka pirka ponno ponno cep ukoyki an=e-hci*
 INDEF Ainu PL always too be.well little little fish catch INDEF.A=eat-PL

‘We Ainu are well as ever; we catch a little fish and eat it.’ — Tangiku & Ogihara 2001: Letter 1; East Sakhalin, Otasan (Sentoku Tarōji)

In the same letter Sentoku writes *an=omante* ‘I send (it)’ of himself alone and addresses Piłsudski as *an=koro nispā* ‘my master’. The indefinite is also the classic narrative first person of folklore: the protagonist of a tale, often a god or hero, is cross-referenced by AN-forms throughout, a logophoric device that keeps the narrator formally distinct from the character whose words the whole tale constitutes (Dal Corso 2025a: 198–199). Piłsudski himself glossed such forms simply as first person singular in his notes (e.g. *sapan* ‘I went down’, (Piłsudski 1912: 37–39)):

- (6) *an-koro nay-he or-ne kimoyki kusu makap-an*
 INDEF.A-have river-POSS place-to hunt PURP go.uphill.PL-INDEF.S

‘I went up along the river I own to hunt.’ — Pilsudski 1912: 133; Dal Corso 2025a: 196; East Sakhalin
 The hero of the tale narrates his own deeds in the indefinite person.

- (7) *Танъ котанъ охтá утáса—анъ кусú áреги анъ.*

tan kotan ohta utasa-an kusu ariki-an
 this village in visit-INDEF.S PURP come.PL-INDEF.S

‘I have come to visit this village.’ — Dobrotvorskij 1875; Sakaguchi 2021; West Sakhalin, West coast, central (recorded 1867–1872)
 A single speaker; ariki is the plural stem of eh ‘come’.

- (8) *kamuy oro-wa mah kayki an-e-kor-te-si kusu*
 god place-from wife even INDEF.A-2SG.O-have-CAUS-PL because

‘since you have been given even a wife by the gods’ — Pilsudski 1912: 23; Dal Corso 2025a: 197; East Sakhalin
 Impersonal passive: an- indexes the backgrounded agent, the recipient ‘you’ is indexed by e-, and the demoted agent kamuy appears as an oblique.

- (9) *anoka kayki mawa-an kusu okay-an*
 INDEF even be.hungry-INDEF.S because exist.PL-INDEF.S

‘I too am here because I am hungry.’ — Murasaki 2001: text 4; West Sakhalin, Odasu (Asai Take)
 A monster speaks of itself in the indefinite; okay is the plural stem of an ‘exist’.

The indefinite dominates Sakhalin first-person text frequency. In the three Sentoku letters we count 23 tokens of *an=* against 17 of *ci=*, 19 of *e=*, and not a single *ku=*; in Yamabe Yasunosuke's (山辺安之助) book-length autobiography the AN-series carries the whole narration, the KU-series appearing once and the CI-series twice (Sakaguchi 2024: 46). The corpus texts recorded by Dobrotvorskij in 1867–1872 use fourth-person narration as their convention, while Asai Take's twentieth-century *tuytah* ‘folktale’ are narrated in the third person with the reportative *manu* and use *ku-* freely in dialogue — so the choice between *ku-* and the indefinite is a matter of register and genre, not of grammar. Within quoted speech both occur, subject to the personal agreement rule of §13.1.

13.3 Object marking

Overt object prefixes exist only for first person and the indefinite; second person objects are indexed by the same *e-/eci-* as subjects, and third person objects are zero. The first person object form divides the island: the west coast has *en-*, as in Hokkaidō, while the east coast generalises *in-* for both the KU- and the CI-series (Satō (1985); Sakaguchi (2024: 33–39)). Earlier work treated *i-* and *in-* as free variants (Chiri (1942: §37)), but Satō showed from the personal agreement rule that *in-* belongs with the KU/CI-series and *i-* with the AN-series (Satō 1985). For the exclusive object, Hattori recorded west-coast *un-*, cognate with Hokkaidō *un-*, though even his Rayciska consultants used the CI-series only rarely (Hattori 1961: 16); Dal Corso notes that east-coast *in-* cannot be derived from it by any regular correspondence, so its origin remains open (Dal Corso 2025a: 205).

- (10) *nee kina e-taa ciki neh en-ee-re*
 some plant 2SG.A-pick if some 1SG.O-eat-CAUS

‘If you pick some butterbur, give me some to eat.’ — Murasaki 2001: text 25; West Sakhalin, Odasu (Asai Take)

- (11) *Самбаку квани охта тамбаку инъ контэ.*

Sambaku kuani ohta tampaku in-konte
 Sambaku 1SG at tobacco 1SG.O-give

‘Sambaku gave me tobacco.’ — Dobrotvorskij 1875: 238; Sakaguchi 2024: ex. 3-8; West Sakhalin, recorded 1867–1872
 The third-person subject is zero; the dative phrase *kuani ohta* ‘to me’ overtly doubles the 1SG object prefix *in-*, recapitulating the recipient (so the KU-series pronoun co-occurs with *in-*). The literal sense is ‘Sambaku gave tobacco to me’.

Sentoku likewise writes *in=konte* ‘give me’ in his third letter (Tangiku & Ogihara 2001). The indefinite object *i-* appears wherever an AN-series referent is acted on, including the narrative first person:

- (12) *ruype kuma ohta oman te ceh huymampa anokay i-ranke*
frozen.fish pole to go.SG and fish inspect INDEF INDEF.O-take.down

‘She went over to the frozen-fish pole, looked the fish over, and took me down.’ — Pilsudski 1912: text 17, ll. 18–20; Sakaguchi 2024: ex. 3-10; East Sakhalin

The narrator has been turned into a fish; the pronoun *anokay* confirms the AN-series.

This prefix is formally identical to the antipassive *i-* (Chapter 14), from which it historically derives (Dal Corso 2025a: 196); synchronically the two are kept apart by the referential (first-person or logophoric) reading of the agreement use. The east coast also has a dedicated first-person passive prefix *inci-*, used when a definite agent acts on the speaker; it serves both the KU- and the CI-series, and corresponds to *enci-/unci-* of north-eastern Hokkaidō (Sakaguchi 2024: 35–37):

- (13) *cookay inci-wenkara ciki hannahkusu pirika kumpe ka han nee nanko*
1PL.EXCL 1.O.PASS-beat if never be.good NMLZ even NEG COP probably

‘If I am beaten, nothing good will ever come of it.’ — Pilsudski 1912: text 19, ll. 23–24; Sakaguchi 2024: ex. 3-13; East Sakhalin

The bear god speaks of himself; *cookay* is the CI-series pronoun with singular reference.

13.4 First–second person interaction forms

When first and second person meet on one transitive verb, the marking is not the simple sum of an A-prefix and an O-prefix: no Ainu variety allows the expected **ku-e-*, **ci-e-*, **ku-eci-* or **e-en-* (Sakaguchi 2024: 32). In Sakhalin the verb has in effect a single prefixal person slot; the object is what gets expressed there, and the subject is either recoverable from context or indexed by a suffix (Dal Corso 2025a: 207–208). Table 2 gives the attested forms; KU- and CI-series behave identically throughout and are collapsed as ‘1’ (Sakaguchi (2024: 53–55); Murasaki (1979: 50); Hattori (1961: 14–16)).

A → O	2SG.O	2PL.O	1.O	INDEF.O
1.A (KU/CI)	<i>eci-</i>	<i>eci-...-yan</i>	—	—
INDEF.A (AN)	<i>an-e-</i>	<i>eci-...-(y)an</i>	—	—
2SG.A	—	—	<i>in-</i> (E) ~ <i>en-</i> (W)	<i>e-i-</i>
2PL.A	—	—	<i>in-...-an</i> (E) ~ <i>en-...-yan</i> (W)	<i>eci-i-</i> (E) ~ <i>i-...-yan</i> (W)

Table 30: First–second person interaction marking

The portmanteau prefixes are best seen in a minimal pair that Dobrotvorskij recorded from nineteenth-century speech — identical word order, the prefix alone carrying the directionality:

- (14) *Чокай эани-еци койки.*

cookay eani eci-koyki
1PL.EXCL 2SG 1SG.A>2SG.O-beat

‘I beat you.’ — Dobrotvorskij 1875: 453; Sakaguchi 2024: ex. 3-31; West Sakhalin, recorded 1867–1872

- (15) *Чокай эани-инь койки.*

cookay eani in-koyki
1PL.EXCL 2SG 2SG.A>1SG.O-beat

‘You beat me.’ — Dobrotvorskij 1875: 453; Sakaguchi 2024: ex. 3-19; West Sakhalin, recorded 1867–1872

The 1→2SG form *eci-* is segmentally identical to the 2PL marker but is a distinct portmanteau; Sentoku uses it to his single addressee:

- (16) *suy yoponi uepekere an ciki eci=nu-re kusu iki*
again later tale exist.SG if 1SG.A>2SG.O=hear-CAUS PURP do

‘If later there is another tale, I will let you hear it.’ — Tangiku & Ogihara 2001: Letter 2; East Sakhalin, Otasan (Sentoku Tarōji)

The AN-series, by contrast, combines its prefixes compositionally: *an-e-* for INDEF→2SG and *e-i-* for 2SG→INDEF. Kindaichi reported precisely this “strict” combinability as the hallmark of Sakhalin (Sakaguchi

2024: 32), though his tables also contain regularised forms (*e-in-*, *eci-in-*, *an-eci-*) for which no textual attestation has been found (Sakaguchi 2024: 38, 51) — a caution against filling paradigms by analogy.

- (17) *anoka* *orowano* *an-e-makan-ke*
INDEF from INDEF.A-2SG.O-go.uphill.SG-CAUS

‘It was I who sent you up into the hills.’ — Pilsudski 1912: text 22, ll. 55–56; Sakaguchi 2024: ex. 3-33; East Sakhalin
The famine god speaks; the pronoun *anoka* confirms the AN-series.

- (18) *sonno* *u-rayki-* *kusuneyke eani* *e-ray* *anoka* *ne* *yahka* *e-i-rayki* *kun* *pe* *ne*
truly *an* RECP-kill- if 2SG 2SG.S-die INDEF COP although 2SG.A- going.to NMLZ COP
INDEF.S INDEF.O-kill

‘If we truly fight, you will die; and me too — you will kill me.’ — Pilsudski 1912: text 12, ll. 150–151; Sakaguchi 2024: ex. 3-20; East Sakhalin

One tale recorded twice on the east coast supplies the same insult in both series, the neatest demonstration of the personal agreement rule: KU-series subject with *in-*, AN-series subject with *e-i-*, never crossed (Sakaguchi 2024: 38):

- (19) ホシキ コキクンペ イネトッパルシ
hosiki *ku-ki* *kun* *pe* *in-etohparusi*
first 1SG.A-do going.to NMLZ 2SG.A>1SG.O-forestall

‘You jumped in ahead of what I was about to do!’ — Urata 1998: 118; Sakaguchi 2024: ex. 3-17; East Sakhalin

- (20) ポシキ アンキクンベ エエトッパルシ
hosiki *an-ki* *kun* *pe* *e-i-etohparusi*
first INDEF.A-do going.to NMLZ 2SG.A-INDEF.O-forestall

‘You jumped in ahead of what I was about to do!’ — Urata 1998: 118–119; Sakaguchi 2024: ex. 3-18; East Sakhalin

Where the second person is plural, the suffix *-(y)an* is added to the portmanteau: *eci-...-yan* for 1→2PL, *in-...-an* (west coast *en-...-yan*, Hattori (1961: 16)) for 2PL→1. Murasaki stated the generalisation for Rayciska — *-yan* appears whenever a first person and a plural second person interact, in either direction (Murasaki 1979: 50) — and Sakaguchi confirms it for the east coast: the suffix marks the plurality of the second person, whatever its role (Sakaguchi 2024: 53–54), an analysis Tangiku also adopts for the contemporary description (Tangiku 2022: 358, in (Bugaeva 2022)). It is surely related to the plural imperative particle *yan* (Murasaki (1979: 30); Chapter 12).

- (21) *nahte* *asip-as* *ciki* *ci-koro-cise-he* *kamuy* *cise* *eci-ko-urenka-re-yan*
then go.out.PL-1PL.EXCL.S if 1PL.EXCL.A-have-house-POSS god house 2PL.O-APPL-be.arranged-CAUS-1.A

‘Then, when I go out, I will make over to you all my house, a god’s house.’ — Pilsudski 1912: text 25, l. 96; Sakaguchi 2024: ex. 3-38; East Sakhalin

- (22) *aynu* *utara* *eci-nee* *kusu* *in-raanup-an* *kusuneyke* *yaykiste*
human PL 2PL.A-COP because 1.O-love-2PL.A if be.dangerous

‘Because you are humans, if you love me, it is dangerous.’ — Pilsudski 1912: text 6, ll. 16–17; Sakaguchi 2024: ex. 3-22; East Sakhalin

A woman warns several men at once; the same verb appears as *in-raanup* with a third-person subject earlier in the tale.

Pilsudski took *inranupan* here to stand “instead of” *inranu* (Pilsudski 1912: 92); the two tokens in the same tale show rather that the suffix tracks the number of the second-person subject (Sakaguchi 2024: 40–41). In the AN-series the same surface string *eci-...-(y)an* is compositional — object *eci-* plus the ordinary indefinite subject suffix *-an* — as shown by tokens co-occurring with the pronoun *anokay* and prefix *an-* in Yamabe’s narrative (Sakaguchi 2024: 49–51). In the corpus, Dobrotvorskij’s conversational texts attest the AN-series 1→2SG combination directly: *an=e=nukara* ‘I met you’ (Dobrotvorskij 1875), (Sakaguchi 2021).

The empty cells of Table 2, finally, are not all gaps of the same standing. The configurations in which agent and object are the same speech-act participant — a first person acting on a first person, a second on a second — can never be filled. Such a clause is necessarily coreferential, and coreference between the two core slots is

not built by stacking an A-prefix on an O-prefix: it is routed through the reflexive prefix *yay-* (marginally *si-*; Chapter 14), which withdraws the object from the person paradigm, so that the derived verb carries a single person marker. East-coast *yay-reske-an* ‘I brought myself up’ shows the reflexive stem taking plain intransitive agreement where *reske* ‘raise’ takes prefixes (Dal Corso 2024: 73–74), and Dobrotvorskij’s conversational texts attest the configuration with the lexicalised *yayko-*:

(23) *Соннока а яйкурамуосьма.*

sonnoka an-yayko-ramuosma
truly INDEF.A-REFL.APPL-be.convinced

‘Truly, we are persuaded.’ — Dobrotvorskij 1875; Sakaguchi 2021: 51; West Sakhalin, West coast, central (recorded 1867–1872)

A notional ‘we convince ourselves (of it)’; the coreferential participant is absorbed by *yay-* (here within lexicalised *yayko-*), and the verb hosts only the single AN-series marker — no 1.A>1.O cell is assembled.

The 1→1 (and 2→2) dashes are thus theorems of the grammar, not accidents of the record, and no portmanteau should be reconstructed for them: reflexivisation empties this corner of the person paradigm in head-marking languages quite generally. The cells pairing the AN-series with the KU- or CI-series follow in turn from the personal agreement rule of §13.1, which bars the two series from a single clause; where a genuinely indefinite agent acts on the speaker — the one non-coreferential reading — the east coast resorts to the passive *inci-* (§13.3). Only what remains once these principles have done their work counts as a documentary gap in the strict sense, and in a closed corpus it must stay one.

13.5 Morphological alignment

Read off the paradigm, the synchronic alignment of Sakhalin Ainu person marking is split by person, and we state it as follows. First person singular is nominative–accusative: one form for S and A (*ku-*), another for O (*en-/in-*). The first person plural exclusive and the indefinite distinguish all three roles — *-as/ci/in-* and *-an/an-/i-* — a tripartite arrangement, with the S-form suffixed and the A-form prefixed. Second persons are neutral (one invariant form each), and third person is neutral by zero. This is Sakaguchi’s classification of the east-coast system (Sakaguchi 2024: 31), and it holds for the west coast with the lexical substitutions noted above.

Dal Corso (2025a) frames the same facts differently. Treating *-as/ci-* and *-an/an-* each as a single marker whose suffixal and prefixal realisations continue one historical item — Tamura’s old observation that *-an* behaves prosodically like the verb *an* ‘exist’ it descends from, reported and extended at Dal Corso (2025a: 190–191) — he dispenses with the tripartite component and describes the basic system as mixed nominative–neutral: nominative for first person and the indefinite, neutral for second and third. The difference is one of segmentation, not of substance: once the distribution (suffix on intransitives, prefix on transitives, distinct O-forms) is stated, nothing empirical hangs on whether S- and A-forms are one morpheme or two. We use the role-explicit glosses INDEF.S, INDEF.A, INDEF.O precisely so that the description does not prejudice this.

The substantive part of Dal Corso’s proposal concerns the interaction forms of §13.4 and their history. He reconstructs Proto Sakhalin Ainu with only *ku-*, *en-* and *e-*, plus a collective verb **asti* that grammaticalised into the plural markers and, fused with the person prefixes, yielded *ci-*, *-as* and *eci-* (Dal Corso 2025a: 203–205). A single prefix slot, accent shift and syllable weight then conspired against double prefixation in first–second person contexts, leaving object-only portmanteaux (*eci-*, *en-/in-*) with subject syncretism (Dal Corso 2025a: 207–208). Finally, the indefinite — already used for a low-referentiality agent acting on a more specific patient, as in impersonal passives — was recruited to disambiguate subject number. The pivot of this last step lies outside the speech-act domain altogether, in 3→3 clauses whose agent is low in referentiality. There *an-* already does inverse-like work: the patient outranks the unexpressed agent, and the clause hovers between an indefinite-agent and a passive reading:

(24) *ahci an-nukara*
old.woman INDEF.A-see

‘Someone sees the old woman; the old woman is seen.’ — Dal Corso 2025a: 210; Sakhalin (dialect not further specified)

On Dal Corso’s analysis *an-* is here an incipient inverse marker: it signals that the agent ranks below the patient, whichever way the clause is translated.

Extending this device into the first–second person domain gave the discontinuous *eci-...-(y)an*, *en-...-yan* and *in-...-an* of §13.4: the prefix goes on indexing the object, while the suffix, in origin the mark of a backgrounded agent, comes to spell out the subject whose person the portmanteau leaves syncretic. Dal Corso reads the result as an emerging, non-prototypical *inverse* — special marking wherever the object outranks the subject in discourse salience — and draws a typological parallel with Algonquian and Athabaskan, where impersonal-agent forms (Meskwaki *-eko-*, Potawatomi *-iko-*) were likewise regrammaticalised as inverse markers (Dal Corso 2025a: 209–213). He is explicit that the system is not uniformly inverse — 1→2SG marking is hierarchical rather than inverse, and 2→1 forms encode object role transparently — and that most reconstructed stages are unattested (Dal Corso 2025a: 202, 212–213).

As a diachronic account we find this persuasive: it derives the portmanteaux from independently motivated phonology, connects the interaction suffix to the indefinite, and explains why the object is always the overtly indexed argument. Synchronically, however, we do not describe Sakhalin Ainu as inverse-aligned. The decisive datum is Sakaguchi's demonstration that in the KU/CI-series the suffix *-(y)an* co-varies with the *number of the second person* in both directions (1→2PL and 2PL→1), while in the AN-series the same string is the ordinary indefinite subject suffix (Sakaguchi 2024: 53–55). A marker whose function is second-person plurality is not an inverse marker. The two readings are nonetheless not flatly incompatible: they approach the same closed corpus from opposite ends, Sakaguchi from the synchronic distribution of the attested cells, Dal Corso from the trajectory that produced them, and the elicitation that would force a choice can no longer be carried out. Where the inverse analysis has genuine synchronic purchase is in the 3→3 indefinite-agent clauses above — a live textual pattern, not a reconstruction — and a reader who follows Dal Corso (2025a) may prefer to gloss such tokens of *an-* as inverse where we keep INDEF. Our synchronic statement is therefore: core alignment nominative–accusative for first person and the indefinite (with role-split tripartite exponence in the CI- and AN-series), neutral for second and third persons; and, overlaid on it, a closed set of portmanteau prefixes and circumfixes for first–second person interactions in which the object controls the prefix slot — a hierarchical trait whose inverse-like asymmetries are the fossil of the development Dal Corso reconstructs.

13.6 Person marking on nouns

The transitive A-prefixes double as possessor indexes on nouns, which then take the possessive (“belonging”) suffix — *ku-matak-ih* ‘my younger sister’ in the following, with the 1→2 portmanteau on the verbs beside it:

- (25) *ku-matak-ih* *eski* *eci-sam-te* *kusu* *eci-ko-ruura* *ruuhe* *nee*
 1SG.A-younger.sister-POSS um 1SG.A>2SG.O-marry-CAUS PURP 1SG.A>2SG.O-APPL-bring COMP COP

‘It was in order to wed my younger sister to you that I brought her to you.’ — Pilsudski 1912: text 25, ll. 82–85; Sakaguchi 2024: ex. 3-28; East Sakhalin

Alongside direct prefixation, possession is freely expressed by a relative construction with the verb *koro* ‘have’: Sentoku's *an=koro nispa* ‘my master’ and *e=koro kotan* ‘your village’, the hero's *an-koro nay* ‘the river I own’ in §13.2, and the compounded *ci-koro-cise-he* ‘the house I own’ in §13.4. Chiri adds that dative (recipient) persons are built on the object forms with an applicative: *in-ko-* ‘to me/us’, *e-ko-* ‘to you’ (Chiri 1942: §37). The morphology and semantics of possessive forms, including the nominal plural *-ahcin*, are treated in Chapter 7.

13.7 Person and verbal number

Person marking interlocks with verbal number at two points (Chapter 8, Chapter 12). First, the suffix *-hci* (east-coast realisations also *-(a)hci*, *-si*, Dal Corso (2024: 62–63)) marks plurality or collectivity of a core argument. Murasaki listed it among the person affixes as an optional third person plural (Murasaki 1979: 49–51), but since it co-occurs with overt person markers of any person — Sentoku's *an=e-hci* ‘we eat it’ above, Murasaki's own *eci-itak-ahci* ‘you all speak’ — we treat it with Dal Corso (2025a: 202) and Sakaguchi (2024) as a number marker, not person agreement. With unmarked third persons it is what signals plurality:

- (26) *iineahsuy* *tan* *cise* *un* *monimahpo* *uta* *hemata* *e-hci* *wa* *okay-ahci* *hii*
 hey this house in girl PL what eat-PL and exist.PL-PL Q

‘Say, what do the girls of this house eat?’ — Murasaki 2001: text 4; West Sakhalin, Odasu (Asai Take)

- (27) *pomoshnik starosta 3 en ranke koro-hci manuy*
 assistant village.head 3 yen each have-PL REP

‘The assistant village heads receive three yen each, they say.’ — Tangiku & Ogiwara 2001: Letter 1; East Sakhalin, Otasan (Sentoku Tarōji)

The first two words are Russian loans; -hci indexes the plural third-person subject.

The controller need not be the subject: Piłsudski already noted of *an-kuyra-si* ‘I stole up on them’ that the plural sign is there “not on account of the subject (‘I’) but because the object (the gods) is in the pl.” (Piłsudski 1912: 40–41). Second, suppletive stem pairs (*oman* ‘go (sg)’ ~ *paye* ‘go (pl)’, *an* ‘exist (sg)’ ~ *okay* ‘exist (pl)’) agree with the grammatical number of the person series: the CI- and AN-series select plural stems even for a single referent — *asip-as*, *paye-as*, *makap-an*, *ariki-an*, *okay-an* in the examples above — while the KU-series takes singular stems (*ku-san*, *e=oman*) (Sakaguchi 2024: 27–29). This grammatical plurality is the reason we segment the indefinite and the CI-series as plural categories whatever they refer to.

Chapter 14 · Valency-changing morphology

Causatives, applicatives, the antipassive, reflexive and reciprocal prefixes, and noun incorporation.

Since every Sakhalin Ainu verb stem is lexically intransitive or transitive (Chapter 12), all adjustment of valency is done by derivational morphology, and the language maintains a full double set of it. Valency is increased by the causative suffixes *-te/-re*, *-ke*, *-ka* and *-yara* and by the applicative prefixes *e-*, *ko-*, *o-*; it is decreased by the antipassive *i-*, the reflexives *yay-* and *si-*, the reciprocal *u-*, and noun incorporation. Chiri catalogued both sets for Sakhalin in one programmatic paragraph (Chiri 1942: §33), and Dal Corso confirms the same inventory in Piłsudski’s east-coast corpus (Dal Corso 2024: 70). The richness and mutual combinability of these ‘voices’ is the celebrated typological trademark of Ainu ((Shibatani 1990); (Bugueva 2022)); this chapter documents what of it is attested in Sakhalin sources, ending with the combinations and a productivity assessment. Two further markers touch valency only obliquely: the resultative *ci-*, which detransitivises by deleting the agent (*ci-ranke* ‘be lowered’ (Chiri 1942: §49), (Dal Corso 2024: 67)), and the impersonal passive built on the indefinite person, treated with person marking in Chapter 13.

Each operation below also has a fixed effect on Chiri’s classification of predicates into complete (object-less) and incomplete (object-requiring) verbs (Chapter 6). The increasers — applicatives and causatives — turn a complete verb into an incomplete one, which thereafter takes the prefixal agreement of a transitive; the decreaseers — antipassive, reflexive, reciprocal, incorporation — fill or remove the object slot and return an incomplete verb to the complete class, with suffixal agreement (Chiri 1942: §33). The whole chapter can accordingly be read as the derivational traffic between Chiri’s two classes, each ‘voice’ a one-way crossing.

14.1 Causatives

The productive causative is *-te* ~ *-re*: *-re* after vowel-final stems, *-te* after consonant-final ones (Chiri 1942: §55), (Tangiku 1998: 229–230), (Tangiku 2022). For the verbs in *-rV* the choice follows from the two stem classes established in Chapter 5: a Class I verb (underlying /r/) counts as consonant-final and takes *-te* with the regular shift of *r* to *n* (*kor-te* → *konte* ‘give’), while a Class II verb (lexically vowel-final) takes *-re* like any vowel stem. Citation forms that look identical therefore part ways under the suffix — *nukara* → *nukante* but *kira* → *kirare* — and the causative is in practice the readiest test of a verb’s class membership. Hokkaidō’s third allomorph *-e* after *r* (*kor-e*, *nukar-e*) is replaced in Sakhalin by *-te* (Chiri 1942: §55), so the *-nte* causatives are among the clearest Sakhalin-defining forms in the verbal morphology. Table 1 lists base ~ causative pairs for which both members are attested in Sakhalin sources.

base	causative	attestation of the causative
<i>ee</i> ‘eat’	<i>eere</i> ‘feed sb sth’	(Murasaki 2001: text 5)
<i>nu</i> ‘hear’	<i>nuure</i> ‘let hear, tell’	(Tangiku & Ogihara 2001: Letter 2)
<i>ipe</i> ‘have a meal’	<i>ipere</i> ‘give sb a meal’	(Murasaki 2001: text 4)
<i>mahnuu</i> ‘take a wife’	<i>mahnuure</i> ‘get sb a wife’	(Murasaki 2001: text 34)
<i>tuuri</i> ‘stretch’	<i>tuurire</i> ‘stretch sth out’	(Murasaki 2001: text 18)
<i>pecika</i> ‘ford, cross’	<i>pecikare</i> ‘take sb across’	(Murasaki 2001: text 18)
<i>koro</i> ‘have’	<i>konte</i> ‘give’	(Tangiku & Ogihara 2001: Letter 1); (Dobrotvorsky 1875); (Sakaguchi 2021)
<i>nukara</i> ‘see’	<i>nukante</i> ‘show’	(Murasaki 2001: text 51)
<i>mokoro</i> ‘sleep’	<i>mokonte</i> ‘put to sleep’	(Murasaki 2001: text 4)
<i>oman</i> ‘go (sg)’	<i>omante</i> ‘send’	(Murasaki 2001: text 51); (Tangiku & Ogihara 2001: Letter 1)
<i>asin</i> ‘go out (sg)’	<i>asinke</i> ‘take out’	(Murasaki 2001: text 5)
<i>ahun</i> ‘go in (sg)’	<i>ahunke</i> ‘put in’	(Murasaki 2001: text 52)
<i>makan</i> ‘go up (sg)’	<i>makanke</i> ‘send up’	(Murasaki 2001: text 35)
<i>tuhse</i> ‘jump’	<i>tuhseka</i> ‘send flying, kick’	(Murasaki 1979: 32); (Murasaki 2001: text 10)

Table 31: Attested base ~ causative pairs

The meaning runs the whole causative range from coercion to giving and letting: *konte* ‘cause to have’ is the ordinary verb ‘give’, and Chiri analysed apparent datives as causees on exactly these verbs (Chiri 1942: §68). The causee is the indexed object (*in=konte* ‘give me’, *eci=nu-re* ‘I will let you hear’; Chapter 13).

- (1) *orowa nean ipe-re teh taa orowa suy taa mokon-te manu*
 then that.one ANTIP.eat-CAUS and EMPH then again EMPH sleep-CAUS REP

‘Then she fed him, and then put him to sleep again.’ — Murasaki 2001: text 4; West Sakhalin, Odasu (Asai Take)

- (2) *oman-te-hci taa nukan-te-hci*
 go.SG-CAUS-PL EMPH see-CAUS-PL

‘They would send him over and show him (to the girls).’ — Murasaki 2001: text 51; West Sakhalin, Odasu (Asai Take)

Both causatives take the plural -hci; the singular stem *oman* agrees with the single causee.

- (3) *acahcipo taa nean unkayoh taa pecika-re teh kema-ha tuuri-re*
 old.woman EMPH that monster EMPH cross-CAUS and leg-POSS stretch-CAUS

‘The old woman ferried that monster across: she stretched out her legs.’ — Murasaki 2001: text 18; West Sakhalin, Odasu (Asai Take)

pecika and *tuuri* are Class II (true vowel-final) stems and take -re (Tangiku 1998: 230).

Three minor causative formatives partition the lexicon with -te/-re. (i) -ke pairs with the *n/p* number stems (Chapter 12): singular *ahun-ke* ‘put one in’ against plural *ahup-te* ‘put many in’, so that the suffix choice itself signals object number (Chiri 1942: §§36, 55); the corpus supplies *asinke*, *ahunke*, *makanke* and the east-coast *an-e-makan-ke* ‘I sent you uphill’ ((Piłsudski 1912: text 22); (Sakaguchi 2024: ex. 3-33)). (ii) -ka derives causatives chiefly from property and posture verbs: *tuhse-ka* ‘kick’ above, *hosipi-ka* ‘send back’ (Chiri 1942: §55). (iii) The thematic vowels -a/-i/-u/-e/-o transitive bare roots (*mak-a* ‘open’, *as-i* ‘stand sth up’) (Chiri 1942: §55). Finally -yara is the indefinite-causee causative, ‘have people do’: Chiri’s *e-yara* ‘have (people) eat’ (Chiri 1942: §55), and lexicalised in *sinukarara* ‘show oneself off’ from *si-nukara-yara* ‘have people look at oneself’ (Murasaki 1979), (Tangiku 1998: 229–230).

Stacked causatives of the shape -te-re or -re-te do not occur in the corpus or in the descriptions we control. Double causation is instead expressed by a causative over an already derived stem (*ipe-re* on antipassive *ipe*, *sinukarara* with reflexive plus -yara) or analytically with the light verb *ki*. We take the gap to be real and systematic — the causee slot, once filled, is not reopened — though a form like *nukante-yara* ‘have people show’ would be the expected shape if it were; finding one in unpublished recordings would settle the point.

14.2 Applicatives

The applicative prefixes promote an oblique participant to object, transitivising intransitives and ditransitivising transitives. The three are not interchangeable: the choice tracks the semantic role of the promoted argument (Dal Corso 2024: 71). *e-* covers location, goal, topic ('about'), path, instrument, purpose, cause and accompaniment — Chiri illustrates each for Sakhalin, e.g. *neampe an-e-mina* 'we laughed about it' from *mina* 'laugh', *oya kotan e-oman* 'he went to another village' (Chiri 1942: §§33, 56). *ko-* adds a goal or addressee, typically a person (Chiri 1942: §58); *o-* adds places only (Chiri 1942: §57). For the southern Hokkaidō dialects Bugaeva's corpus study (Bugaeva 2006: 188–193) quantifies the same division of labour — *ko-* the most and *o-* the least frequent, with about a third of all verb lexemes containing an applicative — and the Sakhalin corpus matches that ranking, as set out below.

- (4) *suma kema e-tuhse-ka*
rock foot APPL-jump-CAUS

'He kicked the rock with his foot.' — Murasaki 1979: 32; Dal Corso 2024: 71; West Sakhalin, Rayciska

The instrument *kema* 'foot', expressed obliquely with *ani* 'with' in the unprefixed construction, is promoted to object by *e-* (Dal Corso 2024: 71).

- (5) *rekoro ituyaskara rekoro an=e=ko-yay-irayki*
very feel.affection very INDEF.A=2SG.O=APPL-REFL-be.grateful

'I feel great affection, and I am deeply grateful to you.' — Dobrotvorskij 1875; Sakaguchi 2021; West Sakhalin, West coast, central (recorded 1867–1872)

ko- adds the addressee, indexed by *e=* 'you', to the intransitive *yay-irayki* 'be thankful'.

Pilsudski's corpus also shows *ko-* in a use where no applied object can be recovered, on zero-valent weather predicates; Dal Corso reads this as a relic of a broader high-transitivity function of the prefix (Dal Corso 2024: 73):

- (6) *ko-sistono to-noske ko-oman kane*
APPL-dawn day-middle APPL-go.SG while

'Day had dawned and, when it had gone on to midday, ...' — Pilsudski 1912: 146; Dal Corso 2024: 73; East Sakhalin
ko-oman '(the day) went on to (midday)' is a regular goal applicative; *ko-sistono* 'it dawned' has no retrievable applied object.

The prefix *o-* is marginal in the Sakhalin corpus: we find it only inside lexicalised place-object verbs (Chiri's locative sense, §57), never as a live alternation parallel to the *e-* pair above. We treat it as a closed lexical class in Sakhalin, against its modest productivity in Hokkaidō (Bugaeva 2006: 193). Applicatives also feed the rest of the system: an applied object is incorporable (§14.6), and the applicative slot is what reciprocal and reflexive prefixes fill in combinations like *u-ko-* and *yay-ko-* (§14.7).

14.3 The antipassive *i-*

Prefixed to a transitive verb, *i-* suppresses the object, yielding an intransitive of generic activity: 'the unspecified, the usual thing' is understood. Chiri glosses it 'things' and lists Sakhalin examples in which the conventionalised object has lexicalised: *i-ku* 'drink (i.e. liquor)' from *ku* 'drink', *inkara* 'look (about)' from *i-nukar(a)*, *i-ruska* 'be angry' from *ruska* 'resent sth' (Chiri 1942: §51). The pair *ipe* 'have a meal' beside transitive *ee* 'eat' (from *i-e*) is the showcase: the two verbs even split as complete versus incomplete predicates in Chiri's classification (Chiri 1942: §30). All four are richly attested in the corpus; *inu* 'listen' (*i-nu* 'hear things') joins them in the Rayciska recordings.

- (7) *okay=an teh ikuu=an raapoke ta*
exist.PL=INDEF.S and ANTIP.drink=INDEF.S while in

'I sat at home, and while I was smoking...' — Kitahara et al. 2003; West Sakhalin, Rayciska (Fujiyama Haru)

iku 'drink/smoke (the usual thing)' takes the intransitive person suffix *=an*, confirming detransitivisation.

- (8) *orowa munaynu taa cise kaa-ta rikin teh nay onne inkara-ha neampe*
then grass.man EMPH house top-LOC climb.SG and river to ANTIP.see-POSS when

'Then the grass-man climbed onto the roof, and when he looked out toward the river, ...' — Murasaki 2001: text 10; West Sakhalin, Odasu (Asai Take)

The antipassive output re-enters derivation freely: *ipe-re* ‘give a meal to’ (§14.1) causativises it, and *inkara* behaves as an underlying *r*-stem (*inkan rusuy*, Chapter 12). The prefix is segmentally identical to the indefinite object marker *i-*, from which it historically derives; the synchronic division — referential first-person or logophoric reading for the agreement marker, generic-object reading here — is drawn in Chapter 13. In our corpus counts, productive (non-lexicalised) antipassives are rare: the construction survives chiefly in the conventionalised activity verbs listed above, which the dictionaries already treat as lemmas ((Hattori 1964)).

14.4 Reflexives *yay-* and *si-*

The reflexive *yay-*, from the root ‘self’, intransitivises a transitive verb, identifying agent and patient: *yay-rayki* ‘kill oneself’, *yay-nuyna* ‘hide oneself’ (Chiri 1942: §53). The valency decrease is visible in person marking: east-coast *yay-reske-an* ‘I raised myself’ takes the intransitive suffix where *reske* ‘raise’ takes prefixes (Dal Corso 2024: 73–74).

- (9) *sine monimahpo yay-reske*
one girl REFL-raise

‘A girl lived all alone (lit. raised herself).’ — Murasaki 2001: text 39; West Sakhalin, Odasu (Asai Take)

- (10) *yay-nuyna teh taa nukara manu*
REFL-hide and EMPH watch REP

‘He hid himself and watched.’ — Murasaki 2001: text 37; West Sakhalin, Odasu (Asai Take)

A second pattern leaves valency intact: *yay-* marks the possessor of an incorporated noun, coreferent with the subject — the ‘quasi-possessive reflexive’ (Bugaeva and Kobayashi, in (Bugaeva 2022)). Pilsudski’s *yay-cise-koro* ‘have a house of one’s own’ (Pilsudski 1912: 238), (Dal Corso 2024: 74) recurs verbatim in Asai Take’s tales:

- (11) *tu horokewpo yay-cise-koro-hci okay-ahci manuyke*
two man REFL.POSS-house-have-PL exist.PL-PL and.then

‘The two men lived each keeping a house of his own, and...’ — Murasaki 2001: text 51; West Sakhalin, Odasu (Asai Take)

Lexicalised *yayko-* (reflexive plus applicative) means ‘by oneself, within oneself’ (Chiri 1942: §25); the nineteenth-century west-coast conversations use it in verbs of mental settlement:

- (12) *Соннока а яйкурамуосьма.*

sonnoka an-yayko-ramuosma
truly INDEF.A-REFL.APPL-be.convinced

‘Truly, we are persuaded.’ — Dobrotvorskij 1875; Sakaguchi 2021; West Sakhalin, West coast, central (recorded 1867–1872) *ramuosma* ‘consent’ is itself *ramu* ‘mind’ incorporated into *osma* ‘rush into’.

The parallel prefix *si-* covers the middle range: with plain transitives ‘oneself’ (*si-etaye* ‘draw back’), with causatives ‘get oneself V-ed’ (*si-kasuy-re* ‘get oneself helped’) and ‘feign’ (*si-ray-re* ‘play dead’) (Chiri 1942: §53); Sakhalin *sinukarara* ‘show oneself off’ (§14.1) builds on it. In the corpus *yay-* outnumbers *si-* by an order of magnitude, and only *yay-* shows the possessive-reflexive use — a distributional statement that is our own.

14.5 Reciprocal and sociative *u-*

The prefix *u-* ‘each other’ intransitivises a transitive verb just as the reflexive does (Chiri 1942: §52). Corpus pairs include *sam* ‘marry’ → *u-sam* ‘marry each other’, *nukara* ‘see’ → *u-nukara* ‘meet’ ((Murasaki 2001: text 34)), *koyki* ‘attack, beat’ → *u-koyki* ‘fight one another’:

- (13) *re horokewpo okore mahnuu re monimahpo u-sam-ahci*
three man all wife.take three girl RECP-marry-PL

‘The three men all took wives; the three girls were wedded (to them).’ — Murasaki 2001: text 44; West Sakhalin, Odasu (Asai Take)

mahnuu ‘take a wife’ is denominal: *mah* ‘wife’ with the verbaliser *-nu(u)* ‘get, have’, beside *hoku-nu* ‘take a husband’ (Chiri 1942: §92).

- (14) *cisekoro horokew*^{*tura*} *taa* *u-koyki-u-koyki-oman-* *taa* *cisekoro horokew*^{*uyki*} *teh* *taa* *ociwe*
 house.owman with EMPH *hci* *hci* *tene* and.then EMPH house.owman kill and EMPH throw
 RECP- RECP- fight- fight-
 PL PL

‘He grappled and grappled with the master of the house, killed him, and threw him out.’ — Murasaki 2001: text 50; West Sakhalin, Odasu (Asai Take)

The co-participant is expressed with comitative *tura* while the verb itself is reciprocal — the regular discontinuous construction.

Beside true reciprocity, *u-* has a sociative-collective use that Chiri singles out as characteristically Sakhalin: on intransitives, with or without a causative, it means ‘in numbers, all together’, sometimes amounting to mere plurality — *utara u-mokon-te* ‘the people slept’, *u-ariki-re* ‘they came along’ (Chiri 1942: §52). Sentoku’s letters supply both this and the combination with *ko-*:

- (15) *обісьта увекарі укобакарі анкіхүі*
opista uwekari ukopakari an=ki-hci
 all RECP.gather RECP.APPL.consult INDEF.A=do-PL

‘We will all gather and take counsel together.’ — Tangiku & Ogihara 2001: Letter 1; East Sakhalin, Otasan (Sentoku Tarōji)

uwekari is *u-* plus *e-kari* with regular glide insertion; *ukopakari* is *u-ko-pakari* ‘reckon with one another’.

Lexicalisation has gone furthest in *ukoyki* itself. In the same letter Sentoku writes *cep ukoyki an=e-hci*, which Tangiku and Ogihara translate ‘we catch a little fish and eat it’ (Tangiku & Ogihara 2001: Letter 1) — here *ukoyki* is not ‘fight each other’ but the collective activity verb ‘take (fish), do the fishing’, built on *koyki* ‘attack, catch (game)’; in Letter 3 it serves as a plain noun, *ukoyki porono an* ‘the catch was plentiful’ (Tangiku & Ogihara 2001: Letter 3). The reciprocal, sociative and lexical readings shade into one another along a path the corpus shows still open.

14.6 Noun incorporation

A noun in its bare conceptual form may be compounded into the verb, ceasing to be a syntactic argument (Chiri 1942: §§24–25), (Tangiku 2022). The canonical case incorporates the object of a transitive activity verb:

- (16) *toyaynu wahka-taa kusu san manu*
 earth.man water-draw PURP go.down.SG REP

‘The earth-man went down to draw water.’ — Murasaki 2001: text 10; West Sakhalin, Odasu (Asai Take)

- (17) *eani wahka-taa eani wahka-taa*
 2SG water-draw 2SG water-draw

‘You fetch the water! You fetch the water!’ — Murasaki 2001: text 9; West Sakhalin, Odasu (Asai Take)

- (18) *orowa nean niina kusu kina-taa kusu makan*
 then that.one firewood.gather PURP plant-pick PURP go.uphill.SG

‘Then she went up into the hills to gather firewood and pick butterbur.’ — Murasaki 2001: text 37; West Sakhalin, Odasu (Asai Take)

niina ‘gather firewood’ (nii ‘wood’) is fully lexicalised; *kina-taa* ‘pick plants’ is built like *wahka-taa*.

That these are words and not object–verb phrases is shown by three convergent facts, which we state from our corpus analysis. First, the plural suffix treats the whole as one stem: *cehkoyki-hci* ‘they caught fish’, *ehahtaa-hci* ‘they dug lily roots’, with *-hci* outside noun plus verb. Second, the incorporated noun is non-referential and unmodifiable: a counted or possessed object always stands outside (*ponno ponno cep ukoyki* ‘catch a few fish’, Letter 1). Third, the complex serves as a bare complement of *kusu* ‘in order to’ exactly like a simple verb, as above. The same nouns occur loose with the light verb (*ceh kii, ehah taa* as two words, Chapter 12), so univerbation is a gradient the speakers exploit.

- (19) *niina-hci koh kuru-hci ceh-koyki-hci koh e-hci ehah-taa-hci koh e-hci*
 firewood.gather-PL whenever burn-PL fish-catch-PL whenever eat-PL lily.root-dig-PL whenever eat-PL

‘They would gather firewood and burn it, catch fish and eat, dig lily roots and eat.’ — Murasaki 2001: text 18; West Sakhalin, Odasu (Asai Take)

Which argument incorporates is constrained by grammatical relation. Shibatani's survey of Ainu concluded that incorporation targets objects (including derived, applicative objects) and patient-like intransitive subjects, obliques only after applicative promotion (Shibatani 1988: 210–211); Satō's Hokkaidō counts give object incorporation 85.9 percent, natural-force subject incorporation (*sir-pirka* type) 6.8 percent, with possessor-stranding and transitive-subject incorporation marginal (Satō 2022: ch. 16, in (Bugueva 2022)). The Sakhalin corpus fits: object incorporation as above; subject incorporation in the weather verbs (*siri-pirika*, Chapter 12); and incorporation of a promoted locative object in *cip-oo* 'ride in a boat' (*cip* 'boat' with *o(o)* 'sit in'):

- (20) *mah wa taa sine aynu taa cip-oo wa san manu*
upstream from EMPH one person EMPH boat-ride and go.down.SG REP

'From upstream a man came down riding in a boat.' — Murasaki 2001: text 37; West Sakhalin, Odasu (Asai Take)

Incorporation also interlocks with the other voices: *yay-* rides on an incorporated possessed noun (*yay-cise-koro*, §14.4), and body-part incorporation underlies lexical prefixes like *he-* 'head' and *ho-* 'rear', which Chiri groups with the detransitivisers (*he-pen-u* 'raise one's face') (Chiri 1942: §§33, 41).

14.7 Combinations and productivity

The valency machinery is freely recursive within fixed scope: each prefix construes with everything to its right, so the outermost prefix corresponds to the outermost derivation. Attested Sakhalin stackings include *u-ko-* (reciprocal filling the applicative's object slot: *u-ko-itak* 'talk to each other' (Chiri 1942: §25), *u-ko-pakari* above), *yay-ko-* (§14.4), *ko-yay-* (applicative over a reflexive base: *an=e=ko-yay-irayki*, §14.2), causative over antipassive (*ipe-re*), reflexive plus indefinite causative (*si-nukara-yara*), and applicative over reflexive-applicative:

- (21) *cise-ta paye-hci ike tani taa acahcipo reekoh e-yayko-nopuru manu*
house-LOC go.PL-PL and now EMPH old.woman very APPL-REFL.APPL-be.delighted REP

'They came to the house, and the old woman was overjoyed at it.' — Murasaki 2001: text 51; West Sakhalin, Odasu (Asai Take)

yayko-nopuru 'rejoice (within oneself)' takes a further *e-* introducing the stimulus.

Our productivity assessment from the corpus, text frequencies weighed against type counts, is as follows. Fully productive: causative *-te/-re* (new formations on loanwords and derived stems; no lexical gaps observed) and the reciprocal *u-*. Productive but lexically anchored: *yay-*, applicative *ko-* and *e-* — most tokens belong to recurring lexemes, yet fresh combinations occur in single texts. Closed or fossilised: applicative *o-*, antipassive *i-*, *si-*, the causatives *-ka*, *-ke* (alive only inside the number-paired motion verbs) and *-yara*. Object incorporation sits in between: the pattern is well represented but its productive core is a small set of subsistence-activity verbs. Dal Corso reaches a compatible conclusion from the other direction, observing that the secondary, valency-preserving functions of *e-*, *ko-* and *yay-* common in Piłsudski's 1900s east-coast corpus have all but vanished from the west-coast recordings of the 1960s, an erosion he connects with the general loss of productive verbal morphology under Japanese contact (Dal Corso 2024: 70–71). The system the last speakers kept intact was the core: causative, reciprocal, reflexive, and the lexicalised stock of everything else.

Chapter 15 · Tense, aspect, and modality

The tenseless verb, aspectual periphrasis with *teh an* and *kusu an*, the perfect, and modal predicates.

The Sakhalin Ainu verb carries no tense morphology. Chiri stated the point twice in his Sakhalin-centred grammar: the absence of tense is one of the defining traits of the verb, and temporal relations are carried instead by aspect (Chiri 1942: §26, §47). Murasaki's conversation course repeats it for the modern learner: the bare verb reports an actual event, present or past, and only unrealised events receive marking (Murasaki 1979), (Murasaki 2025: 23). What the language lacks in tense it makes up for in periphrasis: a set of postverbal elements — converb plus existential, purposive plus existential, and a closed class of auxiliary-like particles — expresses progressive, resultative, completive, iterative, prospective and desiderative values, and shades over into mood. Murasaki catalogued eighteen such 助動詞連語 'auxiliary constructions' for Rayciska (Murasaki 1976: 53–64), and Dal Corso (2018: 178–186) sorted them by syntactic behaviour into aspectual strategies,

light verbs, auxiliaries and final particles. This chapter follows the same path: the tenseless core, the aspect periphrases, the perfect, future and intention, desire, ability, and finally commands and exhortations. One structural thread runs through the modal half of the field and is worth naming at the outset: the copula *ne(e)*. Sakhalin habitually closes a finite clause with a nominalized verb plus the copula (Chapter 20), and it is on that host that the future and dubitative markers grammaticalized — which is why several of them, unlike their Hokkaidō counterparts, carry *nee* as their first element (Bugueva 2016: 106–107) (§15.4). The evidential members of the same postverbal field — *manu* and the possessed evidential nouns — have a chapter of their own (Chapter 16), and negation likewise (Chapter 17).

15.1 The tenseless verb

A bare finite verb is read as present or past according to its lexical aspect. Chiri's rule: verbs of durative meaning (*an* 'exist', *koro* 'have', the copula, and the property verbs) translate as presents — his examples are *mici an-koro* 'I have a grandchild' and *encu an-ne* 'I am an Ainu' — while non-durative verbs used alone are past: *oman* is 'he went', not 'he goes', and Chiri took Batchelor to task for translating *ku-kik* as 'I strike' (Chiri 1942: §47). Temporal adverbs do the work that tense does elsewhere. Murasaki's minimal pair makes the point with one and the same verb form:

- (1) *nuuman Kusiro onne ku=oman*
yesterday Kushiro to 1SG.S=go.SG

'Yesterday I went to Kushiro.' — Murasaki 2025: 23; West Sakhalin

- (2) *kestoasinkoh Kusiro onne ku=oman*
every.day Kushiro to 1SG.S=go.SG

'Every day I go to Kushiro.' — Murasaki 2025: 23; West Sakhalin

Identical verb form; the adverb alone shifts the reading from past to habitual present. The third member of Murasaki's paradigm, *simma suy Kusiro onne ku=oman kusu* 'tomorrow I will go to Kushiro again', adds *kusu* for the unrealised event (§15.4).

The attested time adverbs include *tani* 'now' (over a thousand tokens in our Sakhalin corpus; 1,016 in Asai Take's tales alone), *tanto* 'today', *nuuman* 'yesterday', *simma* 'tomorrow', *nisahta* 'in the morning' and *kestoasinkoh* 'every day'; Murasaki's dialogues combine them freely with bare verbs, as in *tanto nisahta ohorono ku=numa* 'this morning I got up late' (Murasaki 2025: 19, 23, 27). Narrative prose behaves the same way: in Asai Take's tales the chains of bare verbs are past by default, and present-time reference inside dialogue is carried by *tani* or by the progressive periphrasis of §15.2.1. Where previous descriptions speak of a 'future tense' they mean the modal periphery of §15.4: the opposition is realis versus irrealis, not past versus non-past.

15.2 Aspect periphrases

The productive aspect markers are phrasal: a converb or purposive form of the lexical verb followed by *an* 'exist', *kara* 'make', *isam* 'not exist' or *hemaka* 'finish'. Although formally biclausal, these sequences behave as single predicates — they relativize and nominalize as a unit, unlike true purpose or reason clauses — which is why Dal Corso calls them aspectual strategies rather than clause chains (Dal Corso 2018: 184). In the interlinear glosses we keep the pieces literal (*kusu* PURP, *teh* 'and', *an* 'exist'); the aspectual reading belongs to the construction as a whole. Table 1 gives the inventory with the Hokkaidō counterparts that Chiri himself aligned (Chiri 1942: §45–46).

Form	Literal source	Value	Hokkaidō counterpart
<i>kusu an</i>	PURP + ‘exist’	progressive/continuative	<i>kor an</i>
<i>kusu kara</i>	PURP + ‘make’	prospective/conative ‘be about to’	<i>kus-ki, kusu ne</i>
<i>teh an</i>	converb + ‘exist’	resultative/stative	<i>wa an</i>
<i>wa isam</i>	converb + ‘not exist’	conclusive ‘end up, be gone’	<i>wa isam</i>
<i>hemaka</i>	‘finish’	completive	<i>wa okere</i>
<i>an ~ ea</i>	‘exist’ (~ ‘sit’?)	perfect/imperfective (§15.3)	<i>a, wa an</i>
<i>ranke; converb koh</i>	—	iterative/habitual	<i>ranke; kor</i>

Table 32: Sakhalin aspect periphrases

15.2.1 Progressive *kusu an* and prospective *kusu kara* *kusu an* presents an event as ongoing at the reference time. Chiri equated it explicitly with Hokkaidō *kor an*: *oman kusu an* = *oman kor an* ‘he is going’ (Chiri 1942: §46). Murasaki’s textbook sentence shows the present-progressive use with *tani*:

- (3) *ahci tani suukawka kusu an*
old.woman now sew PURP exist.SG

‘The old woman is sewing now.’ — Murasaki 1976: 53; Dal Corso 2018: 181; West Sakhalin, Rayciska (Fujiyama Haru)

- (4) *cis ani unkeyoh taa pahkay teh oman kusu an manu*
cry while Unkeyoh EMPH carry.on.back and go.SG PURP exist.SG REP

‘The Unkeyoh ogre was making off with the crying child on its back.’ — Murasaki 2001: text 1; West Sakhalin, Odasu (Asai Take)

Past progressive: the scene is caught in mid-course at the narrative reference time. The reportative *manu* lies outside the aspect periphrasis (Chapter 16).

In conversation the same periphrasis supplies the everyday present progressive — Murasaki’s lesson title *hemata e=kii kusu an?* ‘what are you doing?’ (Murasaki 2025: 29). Beside it stands *kusu kara*, which Chiri identified as the Sakhalin prospective: *ipe-an kusu an-kara ko* ‘when I was about to eat’ (Chiri 1942: §46). With 79 tokens in Asai Take’s tales (our count) it is fully productive, and under negation it yields a frustrated conative, ‘try to but fail’ (§15.6):

- (5) *wahka an-tah kusu an-kara ike*
water INDEF.A-draw PURP INDEF.A-make when

‘when I was about to draw water’ — Chiri 1942: §46; Sakhalin (dialect not further specified)

A classificatory point needs settling here. *kusu kara* and the particle *ea* of §15.3 — the two markers in this domain that Sakhalin shares with no Hokkaidō variety — have sometimes been carried in the descriptive literature side by side as a pair of ‘duratives’, leaving the impression of two constructions competing for one value. The attestations sort them cleanly. Chiri himself ranges *kusu kara* with the prospectives *kus-ki* and *kusu ne*, judging its use the same (やはり同様の用法), and both of his examples are pre-event — the eating, the water-drawing have not yet begun (Chiri 1942: §45–46). Connected prose confirms the value outside elicitation: in Sentoku Tarōji’s letters *oman kusu kara* is rendered ‘intends to go’ and stands parallel to the plain future *oman kusu ne* of §15.4 in the same correspondence:

- (6) *oman kusu kara*
go.SG PURP make

‘He is going to go.’ — Tangiku & Ogihara 2001: Letter 2; East Sakhalin, Otasan (Sentoku Tarōji)

Epistolary prose about real plans; the periphrasis is pre-event, not ongoing.

ea, by contrast, is what Chiri lists under the durative proper (持続態), ‘be in the course of V-ing’ — an intra-event value (Chiri 1942: §45). The two frames thus differ on both axes at once: in their second element, the full verb *kara* ‘make’ (a preparatory activity) against the bound *ea*, and in their orientation to the event nucleus, before it (*kusu kara*) against inside it (*ea*). Classed this way — prospective against durative — the apparent overlap dissolves; *ea* returns in §15.3, where its east-coast use as a perfect is treated.

15.2.2 Resultative *teh an* and conclusive *wa isam teh an* — converb *teh* plus *an* ‘exist’ — denotes the state resulting from a prior event. Murasaki labels it resultative (Murasaki 1976: 53), (Dal Corso 2018: 178); Dal Corso’s re-edition of the Fujiyama texts glosses it as a perfect. It is the Sakhalin counterpart of Hokkaidō *wa an*: Chiri lists *wa an*, *hine an* and *kane an* as the Hokkaidō resultative set (Chiri 1942: §46), the conjunctive *teh* corresponds to Hokkaidō *wa/tek* as the general converb (Chapter 23), and Satō’s study of Chitose *wa an* (Satō 2006: 43–67) finds there the same stative-perfect value, centred on possession and existence, that *teh an* shows in Sakhalin. The equation Sakhalin *teh an* = Hokkaidō *wa an* is thus a correspondence of both form and function. A *kane an* resultative, by contrast, has no Sakhalin reflex: in our corpus *kane* survives only as a simultaneity converb ‘while’ and as the polite-imperative particle of §15.7 (Murasaki 2025: 27), (Chiri 1942: §119). We count 37 tokens of *teh an* in Asai Take’s 54 tales.

- (7) *nukara koh taa suu kaske ke-oo teh an manu*
see ITR EMPH pot top fat-put.in and exist.SG REP

‘When he looked, fatty meat had been put on top of the pot.’ — Murasaki 2001: 96; Tangiku 2022: §5.1; West Sakhalin, Odasu (Asai Take)

The placing happened off-stage; *teh an* asserts the visible resultant state.

The conclusive *wa isam* ‘end up ...-ing, ... and be gone’ (41 tokens in the same tales) is the one periphrasis where Sakhalin keeps the converb *wa*; it adds to the completed event the disappearance or irrecoverability of its theme (Murasaki 1976: 53):

- (8) *orowa nakene ka oman-u wa isam manu*
then where.to even go.SG-EP and not.exist REP

‘Then he had gone off somewhere and was gone.’ — Murasaki 2001: text 6; West Sakhalin, Odasu (Asai Take)

15.2.3 Iterative and habitual Habitual situations are built with the iterative converb *koh* ‘when-ever’ (Hokkaidō *kor*): chains of *V koh V* clauses are the formula with which Asai Take sketches a protagonist’s way of life at the opening of a tale.

- (9) *niina koh kuru ceh-kii koh ee yuh-kii koh ee nah kii manu*
gather.wood ITR burn fish-do ITR eat game-do ITR eat QUOT do REP

‘He would gather firewood and burn it, catch fish and eat them, catch game and eat it — so he lived.’ — Murasaki 2001: text 19; West Sakhalin, Odasu (Asai Take)

The auxiliary *ranke* marks repeated performance of the event, Chiri’s 反復態 ‘iterative aspect’ (Chiri 1942: §118), (Murasaki 1976: 53). It follows the lexical verb:

- (10) *omantene suke ranke taa eere nah kii manu*
then cook ITR EMPH feed QUOT do REP

‘Then she would cook it each time and feed it to him — so she did.’ — Murasaki 2001: text 25; West Sakhalin, Odasu (Asai Take)

ranke after a verb is the iterative auxiliary; the homophonous transitive verb *ranke* ‘take down’ also occurs in the corpus and must be kept apart.

15.2.4 Completive *hemaka hemaka* ‘finish’ is the completive, ‘do completely, finish doing’ (Chiri 1942: §118), (Murasaki 1976: 53). Unlike the converb periphrases it attaches directly, with no linker, and unlike most of the auxiliaries it survives as an independent verb: asked *hemaka?* ‘is it finished?’, Asai Take answers *hemakaa* ‘it is finished’ (text 2). We count 30 tokens across her tales (21 *hemaka*, 7 plural *hemakahci*, plus lengthened and nominalized forms) — frequent enough to be the default completive, typically closing an episode:

- (11) *caru-hu ohta poro toomus an teh ray hemaka manu*
mouth-POSS in be.big fly exist.SG and die finish REP

‘There was a big fly in his mouth, and he had died.’ — Murasaki 2001: text 27; West Sakhalin, Odasu (Asai Take)

With second persons the completive readily combines with the interrogative nominalization of Chapter 16: *tonnosiki ipe e-hemaka-ha?* ‘have you finished your midday meal?’ (Chiri 1942: §40).

15.3 The perfect: postverbal *an* and *ea*

A bare *an* ‘exist’ directly after the verb — no converb — marks what Murasaki calls resultative-perfective (Murasaki 1976: 53): *ahci suukawka an* ‘the old woman has sewn’ contrasts minimally with the progressive *ahci tani suukawka kusu an* above (Dal Corso 2018: 181). Dal Corso’s later work reads the same *an* as an imperfective which yields a continuous interpretation with atelic predicates (Dal Corso 2024: 72–73); the two accounts agree on the surface fact that postverbal *an* situates the reference time inside the post-event or ongoing state, and differ on which class of predicate is basic. The east coast has beside it a form *ea*, unknown to the western corpora: Chiri records it as a durative particle, ‘be in the course of V-ing’ — the intra-event value that, as §15.2.1 showed, keeps it apart from the prospective *kusu kara* (Chiri 1942: §45, §118) — and in Pilsudski’s texts it serves as a perfect, hosting the *-hV* nominalization in questions — proof that verb plus *ea* has fused into a single predicate (Dal Corso 2018: 182). The origin of *ea* is the open question, and it bears directly on the Hokkaidō comparison. Hokkaidō builds its actional perfect on the particle *a* (plural *rok*), which continues the verb ‘sit’; if *ea* contains the same verb, the two perfects are cognate. Dal Corso argues instead that *ea* is *e-an* with the nasal elided, pointing out that it co-occurs with the progressive *kusu an* on atelic predicates (Dal Corso 2024: 71–73) — a distribution that sits more naturally with the existential than with a posture verb. On that analysis the east-coast perfect is a Sakhalin-internal creation from *an*, the same building block as the postverbal *an* just described, and its resemblance to the Hokkaidō *a* perfect is convergence rather than common inheritance.

- (12) *Hemáta ájnu makánte kuća oxta án eáha.*

hemata aynu makan teh kuca ohta an ea-ha
what person go.up.SG and hut in exist.SG PRF-NMLZ

‘What kind of person has come up and is in my hut?’ — Pilsudski 1912: 134; Dal Corso 2018: 182; East Sakhalin
The hunter returns to find smoke rising from his hut; *ea* bears the interrogative nominalization *-ha*.

15.4 Future and intention

Unrealised events are marked by *kusu* and its compounds. The simplest pattern is predicate-final *kusu*, the regular future-intentional of conversation: *Kusiro onne ku=oman kusu* ‘I am going to Kushiro’ (Murasaki 2025: 19, 23); Chiri classed this *kusu* as the volitional mood of the formal noun *kus(u)* (Chiri 1942: §80), and Tangiku glosses it ‘intention’ (Tangiku 2022: §8.5).

- (13) *simma suy Kusiro onne ku=oman kusu*
tomorrow again Kushiro to 1SG.S=go.SG PURP

‘Tomorrow I will go to Kushiro again.’ — Murasaki 2025: 23; West Sakhalin

Fuller periphrases add a verb to *kusu*: *kusu ne* with the copula and *kusu iki* with *iki* ‘do’. Both are the staple future of Sentoku Tarōji’s letters, which, being real correspondence about plans, attest the intentional reading outside elicitation:

- (14) *uma ihoku ciki nuca kotan oman kusu ne manuy*
horse sell if Russian village go.SG PURP COP REP

‘If he sells the horses, he will go to the Russian village, they say.’ — Tangiku & Ogihara 2001: Letter 1; East Sakhalin, Otasan (Sentoku Tarōji)

- (15) *cokay e=an toko pak no paye=an kusu iki*
1PL.EXCL 2SG.S=exist.SG place until ADV go.PL=INDEF.S PURP do

‘I will go all the way to where you are.’ — Tangiku & Ogihara 2001: Letter 1; East Sakhalin, Otasan (Sentoku Tarōji)
cokay is the CI-series pronoun with singular reference; the verb carries the indefinite suffix (Chapter 13).

The formal nouns *kun ~ kuni* build the expectative, Chiri’s 予期法 ‘what is bound or supposed to happen’ (Chiri 1942: §80): *kun pe ne* ‘is going to, is supposed to’ (examples in Chapter 13) and the periphrasis *kun ki*, attested already in Dobrotvorsky’s nineteenth-century west-coast speeches:

- (16) *Маскинъ ханъ неръ-анъ итах, хаманки кунъ ги кусу.*

maskin han neran itah ham an-ki kun ki kusu
too.much NEG what.kind word NEG INDEF.A-do going.to do PURP

‘I will not say any word at all.’ — Dobrotvorskij 1875; Sakaguchi 2021: 51; West Sakhalin, West coast, central (recorded 1867–1872)

Sakaguchi reads the sentence-final *kusu* as intentional ‘intend to’ and the first *han* as a broken-off repetition of the negator *ham*.

kuni also serves as the complementizer of expected events (Chapter 22). Epistemic futures add the dubitative *nanko* ‘probably’ ~ *nankoro*: *ukoramupirika an-ki nanko* ‘I will (probably) make peace’ in the same Dobrotvorskij speech (Dobrotvorskij 1875), (Sakaguchi 2021); Chiri derives *nankoro* from *nan koro* ‘have the face of’ (Chiri 1942: §118).

The shape of the dubitative differs tellingly between Sakhalin and Hokkaidō. Hokkaidō places bare *nankor* after the finite verb; the Sakhalin corpora far more often show *nee nanko(ro)*, copula first, with the lexical verb in the possessive nominalization of §20.3. Bugaeva reads the prefixed copula as the fingerprint of how the modal field arose: the Sakhalin finite clause is habitually a nominalization closed by *ne(e)* — an originally equational, biclausal structure that has tightened into a single complex predicate (Dal Corso (2018: §5.2.2)) — and the modal particles grammaticalized inside that frame, fusing with the copula they followed (Bugaeva 2016: 106–107, 113). The future *kusu ne* above is the same recruitment of the copula at an earlier stage, and the family extends to the affirmative *neeko* ‘surely’, the dubitative *nean*, and the assertive *nankoro pe ne* of Chapter 22. The host also lets mood stack transparently on aspect: in Fujiyama Haru’s *nuu-hci teh an-i-hi nee nanko* ‘she must have heard it’, the dubitative rides on the copula, which rides on the nominalized resultative *teh an* of §15.2.2 (Bugaeva 2022: Sakhalin text (Dal Corso), ex. 25).

- (17) *temana neeraan an-kara an-kii-hi nee nanko nah ramu*
how what.kind INDEF.A-make INDEF.A-do-POSS COP probably QUOT think

‘She thought: how should I do this?’ — Dal Corso 2018: 25 (Murasaki MRA: 25); West Sakhalin, Rayciska

The dubitative takes the whole nominalized clause *an-kii-hi* as its host, closed by the copula *nee*; *nah ramu* frames the inner speech (Chapter 22).

15.5 Desiderative

The desiderative is *rusuy*, in Sakhalin as in Hokkaidō (Chiri 1942: §118), (Murasaki 1976: 53), (Murasaki 2025). It follows the verb directly and is frequent: 92 tokens in Asai Take’s tales (our count). With *koro* ‘have’ it fuses to *kon rusuy* ‘want (a thing)’ — *eani hemata e-kon-rusuy?* ‘what do you want?’ (text 35) — and Murasaki lists the negative counterpart *etunne* ‘not want to’ among the auxiliaries (Murasaki 1976: 53).

- (18) *ci-reske moromahpo kam-ihci ci-ee rusuy kusu*
1PL.EXCL.A-raise girl flesh-POSS 1PL.EXCL.A-eat DES because

‘because we want to eat the flesh of the girl we raised’ — Murasaki 2001: text 36; West Sakhalin, Odasu (Asai Take)
The ogress speaks of herself alone in the CI-series (Chapter 13).

15.6 Potential

Ability is expressed by *easkay* ‘be able’ and inability by *koyaykus* ‘be unable’ (east coast also *eaykah*); all three descend from full verbs and may still take person prefixes, but normally follow the bare lexical verb as auxiliaries (Chiri 1942: §118), (Murasaki 1976: 53). Syntactically they take a nominalized notional verb as complement — the notional verb can host nominal particles like *ka* ‘even’ — so Dal Corso classes them with the light verbs (Dal Corso 2018: 185). In Asai Take’s tales the negative is five times as frequent as the positive: we count 49 tokens of *koyaykus(-ahci)* against 10 of *easkay(-ahci)* — failure drives plots.

- (19) *naa haciko yahkayki ceh-koyki ka easkay*
still be.small although fish-catch even be.able

‘Though he was still small, he could catch fish.’ — Murasaki 2001: text 23; West Sakhalin, Odasu (Asai Take)
The tale continues *niina ka easkay manu* ‘and he could gather firewood too’.

- (20) *neyke tani taa numa koyaykus manu*
and.then now EMPH get.up be.unable REP

‘And now it could not get up.’ — Murasaki 2001: text 39; West Sakhalin, Odasu (Asai Take)

With the prospective of §15.2.1 the negative potential yields the frustrated attempt: *okoyse kusu kara koyaykus* ‘he tried to urinate but could not’ (text 52). One further form deserves a note of caution: Dobrotvorskij’s dialogues contain a single *eska* (*ene eska e-ki kusu ne yay*), formally reminiscent of *easkay*, but Sakaguchi, whose edition we follow, leaves its meaning undetermined (Sakaguchi 2021: 50); we do not build on it.

15.7 Imperative and hortative

The imperative is the bare stem: *oman!* ‘go!’ (Chiri 1942: §26). No suffix creates the command; the particles that accompany imperatives modulate number and politeness. For plural addressees the particle *yan* is added — in Sakhalin, as Chiri stresses, *yan* attaches to plural verb forms and marks plurality of the addressee rather than creating the imperative itself (Chiri 1942: §119); the same *-yan* reappears inside the first–second person interaction paradigm (Chapter 13).

- (21) *tah kayki sohkara ne koro ike mokoro*
this even mat COP have and sleep

‘Make a mat of this and sleep.’ — Murasaki 2001: text 46; West Sakhalin, Odasu (Asai Take)

Two bare-stem imperatives, *koro* and *mokoro*.

- (22) *eci-sapahka kotahma-re-yan*
2PL-head be.stuck-CAUS-IMP.PL

‘Stick your heads back on, all of you!’ — Murasaki 2001: text 5; West Sakhalin, Odasu (Asai Take)

Addressed to several brothers; the next clause, *eci-macihicin ahun kusu ikii* ‘your wives are about to come in’, shows *kusu iki* in its prospective use (§15.4).

Two final particles soften the command: *waa* ‘do (please)’ and *kanne* ‘please, kindly’ — *en=oponi eh waa* ‘come along behind me’, *ponno en=teere kanne* ‘please wait a little’ (Murasaki 2025: 27). Chiri identifies *kane* ~ *kane* as the Sakhalin equivalent of Hokkaidō imperative *hani* (Chiri 1942: §119), and *nankoro* supplies a still more ceremonious second-person command (Chiri 1942: §118). The prohibitive is built with the preverbal negator *hanka* (Chapter 17).

For the hortative ‘let us’ Chiri notes the absence of Hokkaidō *ro* in his Sakhalin data, citing instead *ke mosso paye-an-hcin!* ‘come, let us go quickly!’ (Chiri 1942: §119) — the indefinite suffix *-an* (Chapter 13) on a plural stem, optionally reinforced by *-hci(n)*. Since the indefinite is the ordinary inclusive ‘we’, a first-person-plural command needs no further marking; interjections (*ke(h)* ‘well then!’) and the particle *anaa* supply the invitation: *tani paye=an anaa* ‘come on, let us go now’ (Murasaki 2025: 23).

- (23) *keh anoka tu-aynu-ne sap-an*
well INDEF two-person-COP go.down.PL-INDEF.S

‘Well then, let the two of us go down.’ — Murasaki 2001: text 30; West Sakhalin, Odasu (Asai Take)

Chapter 16 · Evidentiality

The reportative *manu*, inferential and sensory evidential constructions, and their interaction with person.

Sakhalin Ainu grammaticalizes the source of the speaker’s information. The system has received one monographic treatment, Dal Corso’s comparative study of Ainu evidentiality (Dal Corso 2018), which this chapter follows for the Sakhalin analyses while grounding every form in the corpus. Three layers make up the system. At one end stands the personal-knowledge marking *-hV* (with a zero alternant), a verbal use of the possessive suffix that arose from clause nominalization (Chapter 20; Bugaeva 2016: 104, 113). In the middle sit the inferential and sensory constructions built on a possessed evidential noun plus *an* ‘exist’ — *ruwehe an*, *sirihi an*, *humihhi an*, *hawehe an*. At the other end stands the reportative *manu*, an invariable particle at the very edge of the sentence. Dal Corso arranges the whole set on a single scale of source reliability, from *-hV* (most reliable) through the inferentials down to *manu* (least reliable) (Dal Corso 2018: 303). The reportative dwarfs

everything else in text: of all evidential tokens in Dal Corso's Sakhalin corpora, *manu* accounts for 4,998 of about 5,500, or 91% (Dal Corso 2018: 72). Table 1 gives the inventory.

Form	Composition	Value	ES	WS
<i>-hV ~ -Ø</i>	possessive nominalization	personal knowledge	120	201
<i>ruwehe ne(e)</i>	trace-POSS + copula	inferential (assimilated)	11	0
<i>ruwehe an</i>	trace-POSS + 'exist'	inferential	44	34
<i>sirihi an</i>	appearance-POSS + 'exist'	visual	0	17
<i>humih an</i>	sound-POSS + 'exist'	non-visual sensory	22	9
<i>hawehe an</i>	voice-POSS + 'exist'	auditory	13	32
<i>manu(u) ~ manuy</i>	opaque	reportative	1,474	3,524

Table 33: Sakhalin Ainu evidential forms (token counts from Dal Corso (2018: 72))

16.1 The reportative *manu*

manu signals that the speaker holds the information through someone else's words. Its core uses are hearsay, where the verbal source is left unspecified, and quotative, where the original speaker is named or the report accompanies a quotation (Dal Corso 2018: 295–298). Both are at home in conversation:

- (1) *koro kun mah isam manu*
have going.to wife not.exist REP

'He does not yet have a wife, they say.' — Murasaki 1976: 84; Dal Corso 2018: 296; West Sakhalin, Rayciska (Fujiyama Haru)

Hearsay: the bride's family reports what is known of the suitor.

- (2) *"Ē, ájnu án" náx jé manu.*

ee aynu an nah ye manu
yes person exist.SG QUOT say REP

'"Yes, I am a human being," it said.' — Pilsudski 1912: 100; Dal Corso 2018: 23; East Sakhalin

manu over a verb of saying that hosts a quotation; the quotative complementizer *nah* is doing separate work (Chapter 22).

In direct discourse the bare reportative is rare; speakers prefer to reinforce it lexically with *nu* 'hear' 'so I heard' or *nah ye* 'so they said', which spell out the hearsay or quotative reading that *manu* alone leaves implicit (Dal Corso 2018: 297–298):

- (3) *Neja mánka kiren tán pá né pákhe ráj manu, nú manu.*

neya manka kiren tan pa ne pakhe ray manu nu manu
that be.rich Tungus this year COP spring die REP hear REP

'That rich Tungus died this spring, they said — so he heard.' — Pilsudski 1912: 139; Dal Corso 2018: 297; East Sakhalin

Double *manu*: the inner token reports the death, the outer one reports the hearing.

The letters of Sentoku Tarōji show the reportative at work outside folklore, in the east-coast variant *manuy*. We count seventeen tokens of *manuy* (beside one *manu*) across the three letters: writing real news, Sentoku systematically separates what he vouches for from what he passes on.

- (4) *naa iani e=an manuy ci=nu renkayne*
still 2SG 2SG.S=exist.SG REP 1PL.EXCL.A=hear because

'because we heard that you are still there (we rejoiced)' — Tangiku & Ogihara 2001: Letter 1; East Sakhalin, Otasan (Sentoku Tarōji)

Reportative plus the lexical reinforcement *ci=nu* 'we heard', exactly the conversational pattern of the folklore corpora.

- (5) *herohki kotan an manuy*
herring village exist.SG REP

'there is a herring fishery, they say' — Tangiku & Ogihara 2001: Letter 1; East Sakhalin, Otasan (Sentoku Tarōji)

The third use is the one that floods the narrative corpus: *manu* as a marker of traditional knowledge. In tales the original purveyor of the information is either irretrievable or irrelevant; the teller presents the content as community knowledge transmitted by word of mouth, and *manu* in this function never takes a lexical reinforcement (Dal Corso 2018: 296–299). The formula opens nearly every one of Asai Take's tales:

- (6) *Sannupista sine horokewpo an manu*
 Sannupis.LOC one man exist.SG REP

‘In Sannupis there lived a man, they say.’ — Murasaki 2001: text 19; West Sakhalin, Odasu (Asai Take)

- (7) *Ān-kor hénkihi iréske manu*

an-kor henki-hi i-reske manu
 INDEF.A-have grandfather-POSS INDEF.O-raise REP

‘My grandfather brought me up (the story goes)...’ — Pilsudski 1912: 149; Dal Corso 2018: 296; East Sakhalin

Tale incipit: the first-person protagonist is cross-referenced by the indefinite person (§16.7), and even his own upbringing is presented under the traditional-knowledge reportative.

16.2 The grammar of *manu*

Murasaki listed *manu* among her eighteen ‘auxiliary constructions’ (Murasaki 1976: 53–64), but its syntax sets it apart from the aspectual periphrases and auxiliaries of Chapter 15. Dal Corso ran the whole class through five tests — relativization of the construction as one predicate, independent use of the second element, obligatoriness of a syntactic linker, nominal particles on the first element, and nominalization with *-hV* — and *manu* fails them all: a clause with *manu* cannot be relativized or nominalized, *manu* never predicates on its own, nothing links it to its host, and it carries no morphology whatsoever (Dal Corso 2018: 176–187). He concludes that it is a final particle, like the polite *kannē* or emphatic *naa*: an invariable word that closes the sentence and contributes only procedural meaning. It accordingly takes no person marking and co-occurs freely with every person, including the indefinite of narration, and it follows the entire aspectual complex — *kusu an manu*, *teh an manu*, *wa isam manu* are the normal narrative cadences (Chapter 15). Its origin is opaque; the only published speculation derives it from *hum an* ‘there is the sound’, which Dal Corso reports as unverifiable (Dal Corso 2018: 77). The east-coast variant *manuy* and lengthened *manuu* are phonological, not morphological.

One extension is morphological in appearance only: *manuyke*, the reportative fused with the converb *ike* ‘and (then)’ (Chapter 23). It keeps the narrative under report while linking the clause to what follows; with 260 tokens in Asai Take's tales against 2,357 of plain *manu(u)* (our count), it is the regular way to continue a reported chain without closing the sentence.

- (8) *horokewpo utah situ us-ahci manuyke taa*
 man PL ski put.on-PL REP.and EMPH

‘The men put on their skis, they say, and then...’ — Murasaki 2001: text 51; West Sakhalin, Odasu (Asai Take)

16.3 The inferential *ruwehe an*

The inferential is built on the possessed form of *ruu* ‘trace’: a clause followed by *ruwehe* (reduced *ruhe*, *ruuhe*) plus *an* ‘exist’. The construction is an erstwhile possessive construction: the nominalized clause is the possessor and the sensory noun the possessee, so that *re monimahpo an ruwehe an* below is literally ‘the trace of three girls being there existed’. Synchronically the possessed noun has been pseudo-incorporated into *an* — the two form one phonological word, no modifier may intervene except the possessor clause, and neither piece can be displaced — while still bearing nominal morphology (Dal Corso 2018: 147–156). In the typology of Eurasian noun-modifying clauses the pattern is a gapless one: the scope clause relativizes no argument of its own but simply fills the possessor slot of a perception noun, on the template [possessor clause]–[possessed head]–predicate (Bugueva 2017: 220–221), in contrast to the gapped relative clauses of Chapter 21. The development presupposes the Sakhalin extension of possessive marking to clauses described in Chapter 20 (Bugueva 2016: 104–113). The speaker asserts an event he did not witness, on the strength of its perceptible results:

- (9) *re monimahpo an ruwehe an manu*
 three girl exist.SG trace-POSS exist.SG REP

‘Evidently three girls lived there.’ — Murasaki 2001: text 5; West Sakhalin, Odasu (Asai Take)

The hero enters an empty house and reads the inhabitants off their belongings; the whole inference is then folded under the narrative reportative.

- (10) *unci okore us wa isam ruwe-he an manu*
 fire all go.out and not.exist trace-POSS exist.SG REP

‘The fire had evidently gone out completely.’ — Murasaki 2001: 30; Dal Corso 2018: 152; West Sakhalin, Odasu (Asai Take)

The conclusive aspect periphrasis *wa isam* precedes the evidential and stays within its scope (Chapter 15).

Beside *ruwehe an* the east coast has *ruwehe ne(e)* with the copula, unattested in the west (Dal Corso 2018: 72). The copular variant presents the inference as processed and certain — the content has been thought through and brought into the speaker's own dimension — and it is the only inferential that combines with the personal-knowledge forms (§16.6) (Dal Corso 2018: 300–301). What conditions the choice between the two anchors — viewpoint, person and dialect — is analysed from the copula's side in §18.6:

- (11) *Pon náj oxta ifuráje rúhe né.*

pon nay ohta i-huraye ruhe ne
 be.small river in INDEF.O-wash trace.POSS COP

‘It is clear that she washed me in a small river.’ — Pilsudski 1912: 227; Dal Corso 2018: 21; East Sakhalin

ruhe is the reduced form of *ruwehe*. The narrator, a foundling, reconstructs his own infancy; *i-* is the indefinite object prefix referring to the narrator (Chapter 13).

Person interacts with the construction in a way that confirms the clausal-possessor analysis: a first- or second-person subject of the scope verb is never cross-referenced on the evidential noun (Pilsudski's *otanuye e-kii ruwehe an* ‘you must have drawn on the sand’ shows bare *ruwehe*, not **e-ruwehe*), because the possessor is the whole nominalized clause and hence third person (Dal Corso 2018: 154–155). The one piece of agreement the noun does accept is the nominal plural *-hcin*, optionally echoing a plural subject of the scope verb (§16.4).

16.4 The sensory evidentials

Three further nouns enter the same construction, each naming the channel of perception: *siri* ‘appearance’ for what is seen, *hum* ‘sound’ for what is felt or heard as noise, *haw* ‘voice’ for what is heard as speech or cry. The set matches the Hokkaidō indirect evidentials *siran*, *humas*, *hawas*, with which it shares its origin in sensory nouns — Ainu being one of the typologically rare languages whose evidentials are noun-based (Dal Corso 2018: 73, 156), (Bugueva 2016). The choice among the forms follows a sense hierarchy: vision outranks the other channels, so *ruwehe an* (inference from visible traces) and *sirihi an* (direct visual impression) encode more reliable access than *humihhi an* and *hawehhe an* (Dal Corso 2018: 272–295).

- (12) *tuhso neeno an puy ahun sir-ihhi an manu*
 cave as.if exist.SG hole enter.SG appearance-POSS exist.SG REP

‘It looked as though a hole like a cave opened up.’ — Murasaki 1976: 95; Dal Corso 2018: 21; West Sakhalin, Rayciska (Fujiyama Haru)

Visual appearance, with the proviso that looks may deceive.

- (13) *asin-i-hi neampe okaaketa anayne taa cise onne hemata humi-hi an manu*
 go.out-EP- TOP afterwards after.a.while EMPH house from what sound-POSS exist.SG REP
 NMLZ

‘After she had gone out, after a while there came some sound from inside the house.’ — Murasaki 2001: text 18; West Sakhalin, Odasu (Asai Take)

The interrogative *hemata* ‘what’ inside the construction (‘the sound of something’) is the regular formula for an unidentified stimulus: 31 tokens of *hemata hawehe/humihi* in Asai Take's tales (our count).

- (14) *nean cise onne taa hemata yuukara hawe-he an manu*
 that house from EMPH what song voice-POSS exist.SG REP

‘From inside that house came the sound of someone singing.’ — Murasaki 2001: text 19; West Sakhalin, Odasu (Asai Take)

- (15) *ahkapo-ho-hcin* *cis* *hawe-he-hcin* *an* *kusu* *an* *manu*
 younger.brother-POSS-PL cry voice-POSS-PL exist.SG PURP exist.SG REP

‘It sounded as though her younger brothers were crying.’ — Murasaki 2001: 238; Dal Corso 2018: 150; West Sakhalin, Odasu (Asai Take)

The plural -hcin on the evidential noun echoes the plural subject of *cis*; the progressive *kusu an*, though it follows the evidential, scopes over *cis* ‘cry’ (Chapter 15).

The corpus distribution is sharply skewed. In Asai Take’s 54 tales we count 53 tokens of *hawehe an*, 15 of *ruwehe an*, 10 of *humihhi an* and none of *sirihi an*; the visual form is a Rayciska, west-coast resource (17 tokens in the Fujiyama Haru texts, none in Pilsudski’s east-coast corpus by Dal Corso’s count, where instead *ruwehe ne* appears) (Dal Corso 2018: 72). The auditory form earns its frequency from narrative technique: scenes are routinely perceived through walls and doors, as in the last three examples. Beside its sensory use, *hawehe an* shades into hearsay — ‘there is the voice (saying)’ — bridging towards the reportative; the corpus formula *nah yee hawehe an manu* ‘a voice was heard saying so’ (text 19) stacks voice-perception, quotation and report in one sentence.

Two evidentials may co-occur on one predicate. Personal knowledge -*hV* combines only with *ruwehe ne/an* and *sirihi an*, the high-reliability inferentials, and there it retains only its epistemic force; and *manu*, in its traditional-knowledge function, closes constructions of every type, always last (Dal Corso 2018: 300–302):

- (16) *Ćise oxmaxta šine hójnu ikòkajohó né-ruhe an.*

cise *ohmahta* *sine* *hoynu* *i-ko-kayo-ho* *ne* *ruhe* *an*
 house behind one pine.marten INDEF.O-APPL-call-PERS COP trace.POSS exist.SG

‘Clearly it was a pine marten that had called me from behind the house.’ — Pilsudski 1912: 115; Dal Corso 2018: 301; East Sakhalin

Double evidentiality: the personal-knowledge nominalization -ho under the inferential *ruhe an*; the speaker has worked out whose the voice was.

The reportative supplies a possible third layer, chaining onto the outside of whatever stack the sentence has built. In the following Pilsudski sentence a personal-knowledge nominalization sits under the copula and the whole assertion is then relayed by *manu* — three evidential values in strict continuum order, the most reliable innermost:

- (17) *Ānkonúpuru kusu, tani paxno mójre anhi né manu.*

an-ko-nupuru *kusu* *tani* *pahno* *moyre-an-i-hi* *ne* *manu*
 INDEF.A-APPL-find.interesting because now until be.late-INDEF.S-EP-PERS COP REP

‘Because I found it absorbing, I was late until now (the story goes).’ — Pilsudski 1912: 16; East Sakhalin

The narrating protagonist’s lateness, known at first hand (-hi), is asserted with the copula and folded under the traditional-knowledge reportative of the tale.

16.5 Evidentials as relative tense

Sakhalin Ainu has no tense morphology (Chapter 15), and in narrative much of the temporal load falls on the evidentials. The mechanism is relative tense of the Reichenbachian kind. A sensorial construction asserts that the stimulus — trace, sight, sound, voice — is there for the speaker now, so it pins the reference time to speech time; where the event itself sits is then read off the aspectual class of the predicate in its scope (Dal Corso 2018: §7.3.9). Because a trace can outlast the event that left it, *ruwehe an* is the one form that genuinely alternates: a telic scope predicate places the event before its still-perceptible result, yielding a past inference, while an atelic predicate lets event and trace run concurrently, yielding a present one.

- (18) *hemata ka nii kayki kehke wa cokoko wa isam ruwehe an manu*
 what even tree even break and fell and not.exist trace.POSS exist.SG REP

‘Evidently it had broken and felled every last tree, they say.’ — Murasaki 1979: 99; West Sakhalin, Rayciska (Fujiyama Haru)

Telic scope: the breaking and felling, sealed by conclusive *wa isam* (Chapter 15), is over by the time its results are inspected — a past inference.

- (19) *neya i-reske cikaḥ tani sirukunni kusu utohseka ruhe-hcin an*
 that INDEF.O-raise bird now be.dark because sleep trace.POSS-PL exist.SG

‘Those birds that brought me up must be asleep, now that it has grown dark.’ — Pilsudski 1912: 21; East Sakhalin
 Atelic scope: the sleeping still holds while its signs are perceived — a present inference. The plural -hcin echoes the plural subject (§16.3).

The rest of the set does not alternate. With *ruwehe ne(e)* the copula draws the inferred content into the speaker's here and now, collapsing event, reference and speech time and fixing a present reading whatever the telicity of the scope predicate; and since the sight, sound or voice of *sirihi an*, *humihi an* and *hawehe an* exists only while the event produces it, those three default to the present as well (Dal Corso 2018: §7.3.9) (on the two anchors see §18.6).

Dal Corso left one cell open: motion and possession verbs resist a lexical telicity value, so the past/present prediction for *ruwehe an* appears undetermined over exactly the verbs it most often scopes. For motion the west-coast record decides. A *ruu* is the paradigm of a resultant-state object — what a completed displacement leaves behind — and when a tracker reads a blood-trail in Fujiyama Haru's telling, *oman* ‘go’ under *ruuhe an* can only mean that the going is over:

- (20) *kemi-hi sayayse wa oman ruuhe an hi kusu*
 blood-POSS scatter and go.SG trace.POSS exist.SG NMLZ because

‘because there was a trace that it had shed blood and gone off...’ — Kitahara et al. 2003: tuytah 5; West Sakhalin, Rayciska (Fujiyama Haru)

The trace persists into the inference time; the going that deposited it is necessarily complete — a past reading over the very motion verb whose lexical telicity is indeterminate.

We do not take this as evidence that *oman* is lexically telic. Rather the construction coerces its scope: only a bounded, completed motion can deposit a trace, so the trace-based inferential imposes a telic, result-leaving construal on whatever predicate it takes, and the past reading follows from the construction, not from the verb. Possession resists the same diagnostic for a principled reason — a having that has lapsed leaves nothing behind to read — and we find no token of *ruwehe an* over *koro* ‘have’; that sub-case remains genuinely open.

16.6 Personal knowledge, person, questions and negation

The personal-knowledge evidential asserts that the information sits in someone's own knowledge, whatever the original channel of acquisition. Its exponent is the nominalizing *-hV* of Chapter 20 used sentence-finally, beside a zero alternant; Murasaki described the overt form as an assertive or ‘compassionate question’ ending (Murasaki 1979: 72), and Bugaeva showed that all its uses share presupposed, given content (Bugaeva 2016: 110–113).

- (21) *ku=yee-he sunke*
 1SG.A=say-PERS be.false

‘What I say is a lie.’ — Murasaki 1979: 95; Dal Corso 2018: 73; West Sakhalin, Rayciska (Fujiyama Haru)

Person constrains this evidential less than one might expect. Dal Corso's survey finds *-hV* and *-Ø* distributed almost evenly across first, second and third persons and across affirmative and negative sentences; what varies with person is the pragmatics (Dal Corso 2018: 238–241). In declaratives the source of the knowledge is the speaker. In questions it is the addressee: the speaker uses the personal-knowledge form to signal a presupposition that the hearer knows the answer at first hand (Dal Corso 2018: 237, 252–254). Hence the interrogative is this evidential's natural habitat — Chiri already noted that questions prefer the *-hi* form (Chiri 1942: §40):

- (22) *esinnisah-ta e=numa ike hemata e=kii-hii*
 this.morning-LOC 2SG.S=get.up and what 2SG.A=do-PERS

‘What did you do this morning?’ — Murasaki 1976: 3; Bugaeva 2016: 110; West Sakhalin, Rayciska (Fujiyama Haru)

- (23) *poro iso e=nukara ka hanki-hii*
 be.big bear 2SG.A=see even NEG.do-PERS

‘Have you not seen the big bear?’ — Murasaki 1976: 75; Dal Corso 2018: 239; West Sakhalin, Rayciska (Fujiyama Haru)
 Negation and personal knowledge stack in a question: the negative light verb *hanki* (Chapter 17) hosts the evidential nominalization.

Where no answerer is present at all, the same form builds the rhetorical question with the final *hetaneya*: *hemata kusu enan si hura an-i-hi hetaneya?* ‘why on earth is there such a smell of dung?’ (Pilsudski 1912: 160), (Dal Corso 2018: 252–253). The other evidentials behave asymmetrically here. The inferential and sensory constructions occur freely in content questions inside their scope clause but are not themselves questioned; and the reportative is barred from questions altogether — in Asai Take’s 3,852 corpus sentences *manu* never occurs in a genuine interrogative (our count), as expected of a final particle that marks the assertion as relayed.

Under negation the ordering is rigid: negation is built inside the clause and the evidential closes over it, so the report or inference is about a negative state of affairs, never the reverse. The negators of Chapter 17 precede *manu* (five corpus tokens of the type *V ka hankii manu*, our count), and we find no token in which an evidential noun is negated by *isam* or by a preverbal negator — ‘there is no trace that...’ is simply not how the system works. We analyse this as a scope fact: in the attested material evidentiality in Sakhalin Ainu consistently scopes outside polarity, attaching at the illocutionary level, and a counter-evidential reading would require the lexical strategy (e.g. *nukara ka hankii* ‘did not see it’).

- (24) *mokoro yahka ray ka hankii manu*
 sleep although die even NEG.do REP

‘Though he lay down to sleep, he did not die, they say.’ — Murasaki 2001: text 21; West Sakhalin, Odasu (Asai Take)

16.7 Genre, narration and the reportative

Evidentiality is the grammatical backbone of Sakhalin narrative genre. In the *tuytah* (folktales) of Asai Take, 1,848 of 3,852 corpus sentences — 48% — carry *manu(u)*, 2,357 tokens in all, plus 260 of *manuyke* (our counts, (Murasaki 2001)). Whole pages run with the particle closing every sentence: it is no longer reporting an utterance but framing the tale as traditional knowledge. Dal Corso tracks this shift from reportative proper to direct evidential of shared knowledge across dialects and finds it most advanced exactly in Rayciska and Ussoro, where *manu* is hardly ever a plain hearsay marker any more, while east-coast dialects keep the reportative use alive and disambiguate it with lexical reinforcements; the geography of the shift suggests an areal diffusion through adjacent west-coast villages (Dal Corso 2018: 298–299). Throughout the shift the marker itself never changes: one invariable *manu* covers both functions, and only distribution and reinforcement tell them apart. Hokkaidō Ainu makes the same move overtly, exchanging the hearsay predicate *hawas* for *ruwe ne* when narration presents settled knowledge, which throws the Sakhalin invariance into relief (Dal Corso 2018: 298–299). Whether the west-coast state continues an older east-coast one cannot at present be tested: Pilsudski’s turn-of-the-century materials and Murasaki’s of the 1970s were gathered from different speakers, genres and elicitation settings, and the apparent diachrony may be an artefact of the record.

The reportative frame interlocks with a person convention. In folktales the narrator-protagonist is cross-referenced by the indefinite person *an-*, *-an*, *i-* (Chapter 13): the ‘I’ of the tale is not the teller, and the use of indefinite and first-person-plural marking for the speaker is a fixture of folklore (Dal Corso 2024: 68). Pilsudski’s incipit above (§16.1) shows the pattern: *an-kor henki-hi i-reske manu*, indefinite person under the traditional-knowledge reportative. The letters invert both settings at once: Sentoku writes of himself with *an=* and *ci=*, vouching for his news in unmarked clauses, and reserves *manuy* for the items he has at second hand (§16.1). Genre, person marking and evidential choice are thus a single package: indefinite narration under *manu* for inherited tales, first-person assertion with occasional reportatives for lived report. The same division explains why the inferential and sensory forms cluster inside tales at moments of restricted perception — the protagonist hears crying through a wall, finds a cold hearth — while the personal-knowledge forms cluster in dialogue, where speakers negotiate who knows what.

Chapter 17 · Negation

Standard negation with *hannehka*, the prohibitive, negative existence, and the history of the negative system. Negation is one of the areas where Sakhalin Ainu diverges most sharply from the other Ainu varieties, and one of the few that has received a dedicated modern treatment: Dal Corso (2021) surveys the development of the analytic negative constructions, and Dal Corso (2025b) analyses the full system on the basis of 324 negative

tokens drawn from the East- and West-coast corpora. This chapter describes the attested negators and their syntax, the negative existential, prohibitives and other irrealis negation, and the history of the system. The basic facts are these: Sakhalin's productive verbal negators are built on a single negative element *ham* (the negative existential *isam* stands apart); the older pattern, best preserved on the East coast, attaches *ham*= directly to the negated word; and the West-coast dialects of the twentieth-century recordings negate almost exclusively with analytic constructions, above all *hannehka* and the postverbal *hanki* light-verb pattern.

17.1 The negators and their composition

The negative forms attested in the corpus of Sakhalin Ainu are listed below with the morphological analyses of Dal Corso (2025b: 46–47). All of them contain the element *ham*, which Dal Corso analyses as a proclitic *ham*= on the grounds that it is not class-selective – it attaches to verbs, nouns, the copula, adverbs, and particles – and that it is phonologically dependent on its host.

Form	Composition	Distribution
<i>ham</i> =	NEG=	chiefly East coast; in the West only in fixed phrases
<i>hamo</i> ~ <i>hamu</i>	<i>ham</i> + adverbializer <i>-no</i>	both coasts; caritive ‘without doing’ (§17.6)
<i>hanka</i> (<i>yki</i>)	<i>ham</i> = <i>ka</i> (<i>yki</i>) ‘NEG=even’	prohibitive; <i>hanka</i> West, <i>hankayki</i> East
<i>hanki</i>	<i>ham</i> = <i>ki</i> ‘NEG=do’	postverbal negator; the most frequent strategy overall
<i>hanna</i>	<i>ham</i> = <i>na</i> ‘NEG=also’	East coast (Hunup); apprehensive, negative intention
<i>hannah</i> (<i>ka</i>)	<i>ham</i> = <i>nah</i> (<i>-ka</i>) ‘NEG=so(-even)’	East coast; focus negation, conditionals
<i>hanne</i>	<i>ham</i> = <i>ne</i> ‘NEG=COP’	East coast only
<i>hanneh</i>	<i>ham</i> = <i>ne-p</i> ‘NEG=COP-thing’	both coasts
<i>hannehka</i>	<i>ham</i> = <i>ne-p-ka</i> ‘NEG=COP-thing-even’	West coast only; also negative copula

Table 34: Negative forms of Sakhalin Ainu (after Dal Corso 2025b: 46–47)

The proclitic shows the regular nasal assimilation described in Chapter 5: /m/ assimilates in place to a following consonant, so *ham*=*ne* yields *hanne* and *ham*=*ki* yields *hanki*, while before *s* it dissimilates to *y*, *ham*=*suy* ‘not again’ yielding *haysuy* – the same alternations the nasal-final clitic *an*= shows (Dal Corso 2025b: 45–46). Two further observations bear on the clitic analysis. The first is prosodic: *ham*= ordinarily fuses with its host into one phonological word, re-syllabifying it and drawing the stress onto itself (*hánne*, *háysuy*); but Piłsudski's accented spelling *hám utara* for *ham*=*utara* ‘not the people’ points to a pronunciation [‘ha.mu.ta.ra] where the regular stress rules of a fused CV.CV word would give *[ha.‘mu.ta.ra] – a hint that the negator could still carry stress of its own, as befits an erstwhile independent word recently reduced to clisis (Dal Corso 2025b: 45–46). The second is positional: on verbs, *ham*= stands outside the person prefixes – Piłsudski records *hamecinu* ‘you did not listen’, that is *ham*=*eci-nu*, and *hamanrayki* ‘I did not kill’, that is *ham*=*an-rayki* – a position at odds with ordinary verbal prefixes such as the applicative or the reflexive, which follow agreement. Only two tokens in the whole record show the reverse order person–*ham*: *anameyaynu* (*an-ham-e-yay-nu*) ‘I did not pay attention to it’, from Takoye (Majewicz 1998: 278), and ‘*anhancinukahsiri*’ ‘a land I did not know’, from Rayciska. Dal Corso leaves open whether these show *ham*= on its way to becoming a true inflectional prefix or merely a negator frozen inside a lexicalized word form, on the model of Hokkaidō *somoipere* ‘not feed’ (Dal Corso 2025b: 46).

- (1) *ham*=*eci-nu* *yayne*
 NEG=2PL-hear and.then

‘you did not listen [to my warning], and then...’ – Piłsudski 1912: 121; Dal Corso 2025b: 50; East Sakhalin

17.2 Standard negation

Direct procliticization as in (1) is essentially an East-coast pattern: 19 of its 26 verbal tokens come from East-coast texts, and in the West-coast corpus it survives only inside the folklore formula *hancinukahsiri* ‘an unseen

land’ (Dal Corso 2025b: 49–50). Everywhere else, declarative negation is analytic. The dominant construction – about a third of all negative tokens – places the negated verb first, usually followed by *ka* ‘even’ or its emphatic variant *kayki*, and closes the clause with *hanki*, the negated light verb ‘do’ (Chapter 12 on *ki* ‘do’). The notional verb is a zero-nominalized object of *ki*, and the two coasts disagree over where person marking goes. In the older layout, agreement sits on *ki*, the syntactic head, and the notional verb stays bare; in the layout that dominates the West-coast corpus, agreement has migrated onto the nominalized notional verb, leaving *ki* uninflected (Dal Corso 2025b: 50). The next two examples show the contrast:

- (2) *anoka neampe tohse ka ham=an-ki-no yay-nu-no-ka-an*
 1SG TOP sleep.NMLZ even NEG=INDEF.A-do-ADV REFL-listen-well-CAUS-INDEF.S

‘As for me, I kept an ear out [for strange noises] without sleeping.’ – Pilsudski 1912: 184; Dal Corso 2025b: 50; East Sakhalin

Agreement on the light verb *ki*; the notional verb *tohse* ‘sleep’ is bare.

- (3) *aynu kur-i an-nukara ka ham=kii*
 person figure-POSS INDEF.A-see.NMLZ even NEG=do

‘I did not see the man’s figure.’ – Kitahara 2016: 62; Dal Corso 2025b: 50; West Sakhalin

Agreement on the nominalized notional verb; *ki* carries only the negator.

Dal Corso reads the westward migration of agreement onto the notional verb as a sign that the construction was being reanalysed from two predicates into a single predicate with a postverbal negator (Dal Corso 2021); (Dal Corso 2025b: 57–58); agreement marked on both verbs at once occurs exactly once in the corpus (Dal Corso 2025b: 50).

The same light-verb frame admits the unverbated negators of the *hanne* family in place of *hanki* (Dal Corso 2025b: 51–52):

- (4) *nah kara-hci yahka ampene pirika hanneh kii manuu*
 so make-PL although at.all be.good NEG do REP

‘Although they treated her so, she did not get better at all.’ – Murasaki 1976: 58; Dal Corso 2025b: 51; West Sakhalin, Rayciska

- (5) *si-moy moy-a ka hannehka ki*
 REFL-move-VBLZ even NEG do

‘[The monster] did not move.’ – Murasaki 1976: 99; Dal Corso 2025b: 51; West Sakhalin, Rayciska

The same analytic logic appears in the Dobrotvorskyy-era texts of the 1860s–70s, the oldest connected West-coast record, where *ham* negates the light verb in the intentional construction *kun(i) ki kusu* (Chapter 15):

- (6) *maskin han neran itah ham an-ki kun ki kusu*
 too.much NEG what.kind word NEG INDEF.A-do going.to do PURP

‘I will not say any word at all.’ – Dobrotvorskyy 1875; Sakaguchi 2021; West Sakhalin, West coast, central (recorded 1867–1872)

Note the doubled negation *han ... ham* within one clause, a concord pattern not found in the twentieth-century recordings.

Alongside the light-verb pattern, the unverbated forms *hanne* (East coast), *hanneh*, and *hannehka* (West coast) are used preverbally, immediately before the negated verb:

- (7) *tani nuca renkayne aynu utara ta-ene ahkas kayki hanne ki*
 now Russian because.of person PL there-to walk even NEG do

‘Now, because of the Russians, the Ainu do not travel there.’ – Pilsudski 1912: 140; Dal Corso 2025b: 51; East Sakhalin

- (8) *tani ahturi hannehka ko-tasko*
 now globeflower NEG APPL-tie

‘Now she does not tie [the dried fish] to the globeflower stalks.’ – Murasaki 2001: 112; Dal Corso 2025b: 51; West Sakhalin, Odasu (Asai Take)

The corpus offers no evidence for a dedicated interrogative negation: the strategies above appear unchanged in questions (Dal Corso 2025b: 49). Negative indefinite readings – ‘not a single, none at all’ – are built with a numeral or interrogative plus *ka(yki)* ‘even’ under a negated predicate (Chapter 24 on *ka*):

- (9) *repohpe sine-p kayki hanneh kusa*
 seal one-NMLZ even NEG carry

‘He did not bring home a single seal.’ — Pilsudski 1912: 149; Dal Corso 2025b: 51; East Sakhalin

17.3 Negation of the copula

Equative clauses (Chapter 18) are negated on the East coast by procliticizing *ham=* to the copula *ne*, yielding *hanne*:

- (10) *mahpo-ho-cin aynu po kayki ham=ne*
 daughter-POSS-PL person child even NEG=COP

‘...[but] her daughters were not human children.’ — Pilsudski 1912: 59; Dal Corso 2025b: 49; East Sakhalin

In the West-coast recordings this is impossible — *ham=* is no longer productive there — and the grammaticalized form *hannehka* has taken over the function of a negative copula:

- (11) *anihi humpecehpo hannehka*
 3SG puffer.fish NEG.COP

‘He himself is not a puffer-fish.’ — Murasaki 2001: 198; Dal Corso 2025b: 49; West Sakhalin, Odasu (Asai Take)

17.4 Negative existence and the lexical negatives

Existence is not negated with *ham=*. Sakhalin Ainu instead has a closed set of lexically negative stative verbs, each paired with an affirmative verb that is not normally negated analytically (Dal Corso 2025b: 40):

Negative		Affirmative	
<i>erameskari</i> (WS), <i>eramiskari</i> (ES)	‘not know, not understand’	<i>wante</i>	‘know’
<i>etunne</i>	‘not want’	<i>-rusuy</i>	‘want’
<i>koyaykus</i> , <i>eaykah</i>	‘not be able’	<i>easkay</i>	‘be able’
<i>isam</i>	‘not exist’	<i>an</i> (sg), <i>okay</i> (pl)	‘exist’
<i>sak</i>	‘not have’	<i>koro</i>	‘have’

Table 35: Lexical negative verbs (after Dal Corso 2025b: 40)

The negative existential *isam* (East-coast spelling often *isan*) covers ‘there is no X’, ‘X is absent, away’, and — with a possessor — ‘X has no Y’:

- (12) *tani nuca isan renkayne*
 now Russian not.exist because

‘now that the Russians are gone...’ — Tangiku & Ogihara 2001: Letter 1; East Sakhalin, Otasan (Sentoku Tarōji)

- (13) *maci-hi isan teh okaaketa*
 wife-POSS not.exist and afterwards

‘when his wife was away, then...’ — Murasaki 2001: text 52; West Sakhalin, Odasu (Asai Take)

Following another verb linked with *wa*, *isam* has grammaticalized into a conclusive aspect marker, ‘do completely, off, away’ (Dal Corso (2025b: 41); Chapter 15). The pattern is extremely common in Asai Take’s narratives:

- (14) *hekimoh makanu wa isam manu*
 uphill go.up and not.exist REP

‘(the crow) went off up into the hills, it is said.’ — Murasaki 2001: text 5; West Sakhalin, Odasu (Asai Take)

17.5 Prohibitives, apprehensives, and irrealis negation

The prohibitive is formed by placing *hanka* (West coast) or *hankayki* (East coast) before a verb unmarked for person, exactly like the affirmative imperative (Chapter 15); the command may be reinforced by a final particle such as *wa(a)*, *kanne*, or *yan* (Dal Corso 2025b: 53–54). Sakhalin thus builds its prohibitive from the ordinary negative element, whereas the Hokkaidō dialects use the dedicated prohibitive word *iteki* and its variants (Dal Corso 2025b: 54). The source of the Sakhalin marker is typologically remarkable: prohibitives are

usually recruited from verbs of abstaining or stopping, from modals and desire predicates, or from imperative morphology, whereas *hanka(yki)* is the negation of the additive particle *ka(yki)* ‘even’, as the table of forms above records, a path with only scattered parallels elsewhere and one that Dal Corso traces through the negative contrastive-focus use of negated particles (Dal Corso 2025b: 60). The segmentation itself is old: Chiri already parsed the form as negative *ham* plus emphatic *ka* (Chiri 1942: 536).

- (15) *suy taa hanka oman kanne teere wa*
again there PROH go.SG FIN wait FIN

‘Don’t go off again — wait!’ — Murasaki 2001: text 5; West Sakhalin, Odasu (Asai Take)

In the West-coast corpus *hanka* is not confined to commands addressed to an unmarked verb. The corpus shows it with person-marked verbs in directive contexts, where the negated clause functions adverbially; we take this as a bridge between prohibition and general irrealis negation:

- (16) *hemata e-nukara-ha hanka e-siina kanne yee wa*
what 2SG-see-POSS NEG 2SG-hide FIN say FIN

‘Tell us what you saw, without hiding it!’ — Murasaki 2001: text 34; West Sakhalin, Odasu (Asai Take)

The East-coast dialect of Hunup places *hanna* (variant *hana*) before a verb to express apprehension — ‘lest...’ — and, when the verb carries the intentional auxiliary *kusu iki*, negative intention (Dal Corso 2025b: 54–55):

- (17) *hanna mawa-an- ta wen kiren-or-ne an-kana-te an-e-kun-pe ka ham=ne yawa*
NEG starve- this be.poor Tungus-place- INDEF.A-ask- INDEF.A-eat- even NEG=COP FIN
INDEF.S-and to and COND-thing

‘Lest I starve, should I ask this poor Tungus [for food] and [then find out] it is not something I would eat?’ — Pilsudski 1912: 139; Dal Corso 2025b: 54–55; East Sakhalin, Hunup

Apprehensive *hanna*; the matrix clause is itself a negated equative with *ham=ne*.

The same dialect uses *hannah* to negate the condition of a conditional sentence. The protasis closes with the causal-final linker *kusu* and the topic particle *neyke*; *hannah* follows, before a second *kusu* that partially restates the conditional nexus — in effect the negative-focus syntax of the next section turned on a whole clause. Where the apodosis predicate is nominalized and bears the conditional suffix *-kun*, as below, the sentence is counterfactual: the outcome failed to come about because the condition never held (Dal Corso 2025b: 55–56).

- (18) *tan ohacisuytan mosiri- pahno an kusu neyke han- kusu e-ramu-kayki ham=nemanu*
kes nah mo-
kun-
this empty.house spirit world- until exist.SG be- TOP NEG be- *pe*
end cause cause APPL- even NEG=COREP
soul-be.quiet-
COND-
thing

‘If the empty-house spirit had lived until the end of times, that [would] not have been something [people] would have felt safe about.’ — Pilsudski 1912: 110; Dal Corso 2025b: 55–56; East Sakhalin, Hunup

17.6 Caritive and focus negation

The form *hamo ~ hamu* — historically *ham* plus the adverbializer *-no* — negates a bare, non-finite verb that stands in an adverbial relation to the following predicate, ‘without doing’ (Murasaki (1979: 116); (Dal Corso 2025b: 56)). No clause linker appears, and the negated verb carries no person marking; the caritive form itself signals the dependency (Chapter 23).

- (19) *tu to pahno re to pahno ham ipe makap-an*
two day until three day until NEG eat go.uphill.PL-INDEF.S

‘For two, three days I went on uphill without eating.’ — Pilsudski 1912: 123; Dal Corso 2025b: 56; East Sakhalin

The negator here is bare *ham*, not adverbialized *hamo*; the ‘without eating’ reading comes from the adverbial juxtaposition and the absent clause linker, so the caritive slot admits the plain negator as well as *hamo ~ hamu*.

No split form **ham no* is attested, so the segmentation is a reconstruction; it is nonetheless well supported. The *-no* is the ordinary manner adverbializer, and the missing nasal follows from the same lenition seen in

ponno ~ *pono* ‘a little’; the Hokkaidō negator *somo* continues the same Proto-Ainu verb through just such an adverbialized form, **sam-no*, whose final stress explains its vowel; and *hamo* itself keeps first-syllable stress despite its CV.CV shape, the unstressed final vowel rising to give the variant *hamu* – the profile of a stressed negator unverbated with a stressless suffix (Dal Corso 2025b: 44–45, 47). Chiri supplies an inner-Ainu parallel: in Tokachi (Hokkaidō), *sem-ki-no* ‘without doing’ contracts to *sinkino*, the *s*-form negator feeding a caritive adverbial in *-no* by the same route (Chiri 1942: 536).

Finally, the East-coast texts attest *ham=* directly on a noun – always *utara* ‘people’ – on the adverbs *suy* ‘again’ and *ene* ‘like this’, and once on the topic particle *neampe*, marking contrastive focus negation: the marked constituent is singled out as not satisfying the predicate, against some alternative that does – ‘not X (but Y)’ (Dal Corso 2025b: 52–53):

- (20) *an-matak-hi* *nah* *ye* *yahka* *ham=utara* *nu-hci*
 INDEF-younger.sister-POSS so say though NEG=people hear-PL

‘Although my younger sister said so, it was not [you] men who listened to her [but I did].’ – Pilsudski 1912: 120; Dal Corso 2025b: 53; East Sakhalin

17.7 History of the system

Dal Corso reconstructs the whole family of negators from a Proto-Ainu stative transitive verb **sam* ‘not exist in (a place)’ (Dal Corso 2025b: 43–45). One line of development antipassivized the verb with *i-* (Chapter 14), yielding the negative existential *isam* ‘not exist (anywhere)’; the other turned the bare verb into a negative marker placed before the negated element, which weakened to the proclitic *ham=*, with regular **s* > *h* outside the neighbourhood of *i*. The Hokkaidō negator *somo* continues the same verb through an adverbialized form **sam-no*; the Sakhalin caritive *hamo* is its direct cognate, as set out in the previous section.

Dal Corso sets this prehistory within two named typological cycles. The step from negative verb to negative marker instantiates the Givón Cycle, by which erstwhile negative verbs are worn down into negators (Givón 1978). On Croft’s Negative Existential Cycle (Croft 1991), the reconstructed stage in which one form, **sam*, served as both verbal negator and negative existential is Type C – the rarest of Croft’s types, reported otherwise mainly from Dravidian and Polynesian languages (Veselinova 2016) (Dal Corso 2025b: 56–57). The split into existential *isam* and verbal *ham=*, driven by the antipassive and by the sound change, took Sakhalin Ainu directly to Type B, with distinct negators for the two domains, skipping the intermediate Type A in which one strategy covers both; since the stages of the cycle need not be sequential, the skip is unexceptional (Dal Corso 2025b: 57).

From procliticized *ham=* (Stage 1) the analytic system then unfolded in implicational order (Dal Corso 2025b: 57–59). Negation of the copula gave *hanne*, which in the East was reanalysed as a preverbal verbal negator and drawn into the light-verb frame to mark emphasis (Stage 2); because *hanne* was still felt to be copular, the construction was renewed with overt nominalization as *hanneh* (*ham=ne-p*) and emphatic *hannehka* (Stage 3). In the West, where *ham=* had ceased to be productive, preverbal *hanneh(ka)* was then pressed into service as the plain, non-emphatic negator (Stage 4) – an emphatic form bleaching into the default, the profile familiar from the Jespersen cycle of negator renewal. Finally, by ellipsis of *ki*, *hannehka* came to stand as a negative copula in its own right (Stage 5): a verbal negator turning copular with no new affirmative existential beside it, a backward transition in the Negative Existential Cycle that Dal Corso suggests is so far unparalleled (Dal Corso 2025b: 58–59). The light-verb frame itself need not be a purely internal invention: negation through a nominalized predicate governed by a dummy verb recurs in Hokkaidō Ainu, in Nivkh, and in Chukotko-Kamchatkan, and diffusion around the southern Okhotsk Sea is a live possibility (Dal Corso 2021: 223–225); (Dal Corso 2025b: 57–58).

The corpus counts lay the cline out on the ground. Every later stage is the stronger represented the further west and the later the recording: synthetic *ham=* and *hanne* are essentially Eastern, *hanki* and *hannehka* overwhelmingly or exclusively Western (Dal Corso 2025b: 47).

Negator	East coast	West coast
<i>ham=</i> (synthetic)	25 (22.5%)	7 (3.3%; all in the fixed formula)
<i>hanne</i>	42 (37.8%)	0
<i>hanneh</i>	3 (2.7%)	6 (2.8%)
<i>hanki</i>	11 (9.9%)	99 (46.5%)
<i>hannehka</i>	0	56 (26.3%)

Table 36: Tokens of the principal negators by dialect group (after Dal Corso 2025b: 47). Percentages are of all negation tokens in each group (East 111, West 213).

The same drift can be read against time within the western records themselves. Of the negation tokens in Dobrotvorskij's West-coast materials of the 1860s–70s, the preverbal clitic strategy still accounts for 68.4%; in Murasaki's 1983 data it has fallen to 3.7%, while the postverbal light-verb construction has risen from zero to 49% (Dal Corso 2021: 223). Taken at face value, the figures compress the synthetic-to-analytic turnover into roughly forty years, which Dal Corso himself flags as suspiciously fast; more plausibly the two dialect groups moved at different speeds, the East conservative and the West innovating, so that the staged cline is as much a geographic gradient as a chronological sequence – an implicational ordering of overlapping strategies rather than a dated relay (Dal Corso 2025b: 61).

The Hokkaidō negator *somo* itself surfaces in Sakhalin sources only marginally: Dal Corso finds it in Piłsudski's Tunayci texts, told by speakers who had lived in Hokkaidō, and treats it as contact influence (Dal Corso 2025b: 41). The corpus adds one more such setting: in Sentoku Tarōji's letters of 1906 the fixed conditional phrase *somo ne ciki* 'if (it) is not (so)' appears beside native *ham*-negation, fitting the same profile of a literate East-coast writer in close contact with Hokkaidō usage.

- (21) *hospi sonko i=konte somo ne ciki*
 return message 1SG.O=give NEG COP if

'if you do not send me a reply...' – Tangiku & Ogihara 2001: Letter 1; East Sakhalin, Otasan (Sentoku Tarōji)

On the object form *i=* (beside *in=*) see Chapter 13; *somo* here is the Hokkaidō negator in a fixed phrase.

Part V · Syntax

Chapter 18 · Copula, existence, and possession clauses

The copula *ne*, the existential verbs *an* and *okay*, locational clauses, and existential possession.

Two small verbs carry much of the weight of Sakhalin Ainu syntax: the copula *ne*, which builds equative and ascriptive clauses, and the existential *an* with its plural partner *okay*, which build existence, location, and — together with *koro* ‘have’ — possession clauses. Both also live a second life inside grammar: *ne* in the future construction *kusu ne* and in the topic particle *neampe*, *an* in aspectual periphrases and in the evidential constructions. This chapter describes the clause types; the grammaticalized uses are treated in Chapter 15 and Chapter 16.

18.1 Copular clauses

An equative clause has the syntax of a transitive clause: the copular object — the predicate nominal — stands immediately before *ne* and is never case-marked, an overt subject precedes it, and a non-third-person subject is indexed on the copula itself with the transitive subject prefixes (Dal Corso 2025b: 48). In the narrative corpus the copula typically appears with a long vowel, *nee*, and is followed by the reportative *manu* (Chapter 16); on vowel length see Chapter 4.

- (1) *inkara-hci-hi neampe ota nee manu*
look-PL-POSS TOP sand COP REP

‘When they looked — it was sand.’ — Murasaki 2001: text 3; West Sakhalin, Odasu (Asai Take)

The possessive suffix here nominalizes the clause, ‘their looking’; see Chapter 20.

A first- or second-person subject is indexed on the copula itself, as the conversational record makes plain; the predicate nominal may identify the referent by name (an identificational clause) or assign it to a class (a classificatory one), a contrast Murasaki’s conversation course draws in a single exchange:

- (2) *Taroo ku=nee*
Tarō 1SG=COP

‘I am Tarō.’ — Murasaki 2025: lesson 2; West Sakhalin

The overt 1SG prefix *ku=* on the copula, in an identificational clause naming the referent.

- (3) *kuani neyke siisam mahtekuh ku=nee*
1SG TOP Wajin woman 1SG=COP

‘I am a Wajin woman.’ (lit. ‘as for me, I am a Wajin woman’) — Murasaki 2025: lesson 3; West Sakhalin

The same speaker has just identified herself by name (*kuani neyke* Kiyoko *ku=nee* ‘I am Kiyoko’); here the predicate nominal *siisam mahtekuh* assigns her to a class. *neyke* is the conditional-based topic ‘as for’ (Chapter 23).

The predicate nominal may itself be a noun phrase of any internal complexity, including one headed by a relative clause (Chapter 21):

- (4) *kimma tura san monimahpo nee manu*
from.mountains together come.down.SG girl COP REP

‘She was the girl who had come down with him from the mountains.’ — Murasaki 2001: text 43; West Sakhalin, Odasu (Asai Take)

Negation of the copula differs by coast: East-coast *ham=ne*, West-coast *hannehka* as a suppletive negative copula (§17.3).

18.2 Existential clauses

Existence is predicated with the suppletive pair *an* ‘exist’ (singular) and *okay* (plural) – the same verbs that underlie the aspectual periphrases of Chapter 15. Their suppletion is part of the wider system of verbal number described in Chapter 12; the plural stem may carry the plural suffix *-hci* as well. A bare existential clause with an indefinite noun phrase is the standard way of introducing a new referent, and the formula *N sineh an manu* ‘there was one N’ opens tale after tale in the Asai Take corpus (Chapter 24):

- (5) *ahci sineh an manu*
old.woman one exist.SG REP

‘There was an old woman, it is said.’ – Murasaki 2001: text 1; West Sakhalin, Odasu (Asai Take)

- (6) *Sannupis-ta re horokewpo okay-hci*
Sannupis-LOC three young.man exist.PL-PL

‘In Sannupis there lived three young men.’ – Murasaki 2001: text 5; West Sakhalin, Odasu (Asai Take)

The negative counterpart is the lexical negative existential *isam*, never a negated *an* (§17.4).

18.3 Location

A locative existential adds a place expression – a locative-marked noun or a relational noun phrase (Chapter 11) – before the existential verb. With a definite subject the same frame reads as a locational statement, ‘X is in/at Y’:

- (7) *niinum tuh pis ohta an manu*
wooden.ball two piece in.it exist.SG REP

‘Two wooden balls were in it.’ – Murasaki 2001: text 52; West Sakhalin, Odasu (Asai Take)

Singular *an* with a numeral-quantified subject; agreement with conjoined and quantified subjects is variable in the corpus.

Existential and locational clauses freely modify nouns, since any clause may stand before a head noun (Chapter 21); the frame *cise ohta an N* ‘the N who was in the house’ is among the most common relative patterns in the narrative corpus:

- (8) *cise ohta an monimahpo nah kii manu*
house in exist.SG girl QUOT do REP

‘The girl who was in the house said this.’ – Murasaki 2001: text 52; West Sakhalin, Odasu (Asai Take)

18.4 Possession clauses

Predicative possession uses the transitive verb *koro* ‘have’, whose possessor is the A argument and possessee the O argument. With kin and offspring the verb also covers ‘get, come to have’:

- (9) *ohkayo poo sineh mahtekuh poo reh pis ka koro manu*
male child one female child three piece even have REP

‘She had one boy and three girls.’ – Murasaki 2001: text 19; West Sakhalin, Odasu (Asai Take)

The same verb supplies an adnominal possessive construction: a relative clause on *koro*, as in Sentoku Tarōji’s salutation *an=koro nispa* ‘my master’, literally ‘the master whom we have’ ((Tangiku & Ogihara 2001: Letter 1)), beside the morphological possessive of Chapter 17.

7. Negative possession is expressed by the lexical negatives *sak*

‘not have’ or, with the possessee as subject, *isam* ‘not exist’ (§17.4; (Dal Corso 2025b: 40)).

18.5 Grammaticalized offshoots

Both core verbs feed grammar beyond the simple clause. The copula combines with the purposive *kusu* in the intentional future *kusu ne* (Chapter 15), and the ubiquitous topic particle *neampe* is transparently *ne-an-pe* ‘the thing which is’ – a frozen headless relative on the copula and existential (Chapter 24; Chapter 20). The existential *an* closes the stative periphrases *teh an* and *kusu an* (Chapter 15) and the evidential constructions of the type *ruwehe an* (Chapter 16). These uses are the synchronic face of a long history: the person marker *-an* itself continues the verb *an* (Chapter 13; (Dal Corso 2025b: 44)).

18.6 Copula and existential inside the inferentials

The evidential constructions of Chapter 16 end in one of this chapter's two verbs: a possessed perception noun, preceded by its clausal possessor, is anchored either by *an* or by the copula. The two anchors are not interchangeable. The first restriction is lexical: only *ruwehe* 'its trace' admits both (*ruwehe an* ~ *ruwehe ne(e)*), while *sirihi*, *humihi* and *hawehe* take the existential only (Dal Corso 2018: 72).

Where both anchors are available, the choice tracks the speaker's viewpoint on the inferred event. We analyse existential *an* as the external construal: the situation is reconstructed from outside, as a closed whole read off its trace, and the form carries a mirative shading of the speaker's own prior unawareness. The copula construes the situation from inside, as content the speaker has processed and made her own. The distributional signature of the split is person: the copular tokens in Dal Corso's corpus overwhelmingly have a first- or second-person participant inside the scope of the inference, while the existential resists speech-act participants there (Dal Corso 2018: §7.3.5). One infers with the copula about events one was part of; startled inference about wholly external events selects the existential. The opposition thus redeploys the internal/imperfective versus external/perfective contrast of viewpoint aspect, applied not to a lexical verb but to the speaker's stance on an inferred proposition. Its temporal corollary — the copula forces a present reading, the existential leaves telicity to decide between past and present — is set out in §16.5.

The alternation is also dialect-diagnostic. *ruwehe ne(e)* is confined to the east coast, which never developed the visual *sirihi an*; the west coast has *sirihi an* and lacks the copular form (Dal Corso 2018: 72, §7.3.4). The two are in near-complementary distribution across the dialect boundary and do comparable work: each coast opposes a processed, speaker-internal inference to a raw external one, but encodes the processed pole with the copula in the East and with a dedicated perception noun in the West.

Finally, the two verbs can co-occur in one evidential predicate, each in its own slot. In the sentence below the copula equates a nominalized clause with the evidential noun, and the existential then anchors the whole as a present state — stacking, not redundancy:

- (10) *itah* *ku=wante-he* *nee* *ruu-he* *an*
 speech 1SG.A=know-NMLZ COP trace-POSS exist.SG

'I do think I remember the language.' (lit. 'there is the trace of the speech being what I know') — Bugaeva 2016: 29; West Sakhalin

Chapter 19 · The simple clause and word order

Argument expression and omission, constituent order, adjuncts, and clause types including questions and commands.

The Sakhalin Ainu clause is head-marking: the verb indexes its arguments with the person affixes of Chapter 13, and noun phrases appear only when reference needs to be established or shifted. The one extended syntactic description of a Sakhalin variety is Furukawa's generative sketch of Rayciska, based on her own fieldwork with Fujiyama Haru (Furukawa 1967); its rules and examples, read together with the narrative corpus, underpin much of this chapter. Furukawa's clause template — subject, predicate, auxiliary, and an obligatory final-particle slot carried by intonation (Furukawa 1967: 102, 105) — is a fair summary of what the texts show: a verb-final clause closed by a small grammatical tail of auxiliaries, evidentials, and final particles.

19.1 Arguments and their omission

Person marking makes independent pronouns redundant. Furukawa states the distribution plainly: *kuani* 'I', *eani* 'you', *anoka* 'we', *ecioka* 'you (pl)' are used only when the speaker points to the referent itself, while in ordinary clauses the affixes alone carry the relations Furukawa (1967: 102–103) (Chapter 9). Third persons are unmarked on the verb, so a third-person argument that is not recoverable from context is the one that appears as an overt noun phrase:

- (1) *tampe* *ku-nukara*
 this 1SG-see

'I see this.' — Furukawa 1967: 103; West Sakhalin, Rayciska (Fujiyama Haru)

- (2) *taranoka aynu uta tampe en-konte-hci*
 that.PL person PL this 1SG.O-give-PL

‘Those people gave this to me.’ — Furukawa 1967: 117; West Sakhalin, Rayciska (Fujiyama Haru)

Connected narrative pushes the same economy further. In the Asai Take corpus, runs of clauses with no overt noun phrase at all are the norm once protagonists are established, with reference managed by the verbal morphology, by the anaphoric demonstrative *neya* when needed, and by the discourse conventions described in Chapter 24. Overt noun phrases cluster at episode boundaries, where participants enter or change; this distribution is our generalisation over the corpus, and it matches Furukawa's elicited judgements.

19.2 Constituent order

The pragmatically neutral orders are subject–verb and agent–object–verb; obliques and adverbs precede the verb, and nothing follows it except the grammatical tail — auxiliaries, evidentials, and final particles (Furukawa (1967: 102); (Dal Corso 2025b: 39)). Example (2) above shows the full A–O–V layout with a ditransitive verb. A clause with a chained posture predicate illustrates the same head-final profile:

- (3) *ahci kuwa koro kanne ahkas kusu an*
 old.woman stick hold while walk PROG exist.SG

‘The old woman is walking, holding a stick.’ — Furukawa 1967: 119; West Sakhalin, Rayciska (Fujiyama Haru)

On the progressive periphrasis *kusu an* see Chapter 15.

Departures from verb-finality in the spoken corpus are afterthoughts: a noun phrase identifying a referent is added after the predicate (and after the reportative *manu*), with the prosody of a separate tail. We analyse these as right-dislocation rather than basic order; they typically resolve the reference of a quoted speaker:

- (4) *nah yee manu tu horokewpo*
 QUOT say REP two young.man

‘...so said the two young men.’ — Murasaki 2001: text 3; West Sakhalin, Odasu (Asai Take)

The dislocated subject follows the clause-final reportative; the preceding quotation is a prohibitive (§17.5).

19.3 The noun phrase

Within the noun phrase the order is fixed: demonstrative, numeral, modifying clause, head noun, plural word — Furukawa's rule NP → (Det_d) (Det_n) (clause) N (plural) (Furukawa 1967: 102). Modifying ‘adjectives’ are clauses (Chapter 6; Chapter 21), and the plural words *uta(h)* (Chapter 8), *oy*, and *ikiri* close the phrase. The possessive morphology wraps around the whole phrase, not just the head noun: Furukawa records *e-kurasno seta-ha* ‘your black dog’ and, with a numeral,

- (5) *e-tu-seta-ha*
 2SG-two-dog-POSS

‘your two dogs’ — Furukawa 1967: 104; West Sakhalin, Rayciska (Fujiyama Haru)

The 2SG prefix and the possessive suffix enclose numeral plus noun; see Chapter 7.

19.4 Questions

Sakhalin Ainu marks polar questions with intonation and final particles, never with word-order change. The striking fact, recorded by Furukawa for Rayciska, is that the interrogative contour is FALLING while declaratives end high: the same string *tara aynu omanihi* is ‘that person went’ with rising pitch and ‘did that person go?’ with falling pitch (Furukawa 1967: 109). Final particles split by the same contour — *wa* with rising pitch commands, with falling pitch questions (Furukawa 1967: 105):

- (6) *neya nispa eh wa*
 that gentleman come.SG Q

‘Has the gentleman come?’ — Furukawa 1967: 105; West Sakhalin, Rayciska (Fujiyama Haru)

Falling intonation; the same segmental string with rising pitch is an imperative ‘come!’.

- (7) *pirika-ha aa*
be.good-POSS Q

‘Has it healed?’ — Furukawa 1967: 109; West Sakhalin, Rayciska (Fujiyama Haru)

The particle *aa* questions a nominalized (perfect) form in *-hV*; see Chapter 20.

The default polar particle of the modern conversational record is *hee*, which closes copular and verbal questions alike; it is the interrogative particle of Murasaki’s Enciw course (Chapter 9), and the rest of this grammar relies on it. A copular polar question:

- (8) *tara mahtekuh neyke Husako nee hee*
that woman TOP Husako COP Q

‘Is that woman Husako?’ — Murasaki 2025: lesson 3; West Sakhalin

The negative answer in the same exchange is *Husako ka hannehka* ‘she is not Husako’ (§17.3).

Sentoku Tarōji’s letters attest the particle *ya* in the same function, stacked after *hetane* in *temana an okay e=kihi hetane ya* ‘how are you faring?’ (Chapter 26). Content questions leave the interrogative word in its clause-internal position (Chapter 9 for the inventory); no fronting is attested in the corpus:

- (9) *hemata e-nukara haa*
what 2SG-see Q

‘What did you see?’ — Murasaki 2001: text 34; West Sakhalin, Odasu (Asai Take)

19.5 Commands and invitations

The imperative is the bare verb, normally supported by a final particle scaled to politeness: *kane* for polite requests, *wa* for plain commands, *kah* for brusque ones used with intimates (Furukawa 1967: 105–108). The prohibitive counterpart with *hanka* is described in §17.5.

- (10) *tah ampa kah*
this carry FIN

‘Take this!’ — Furukawa 1967: 108; West Sakhalin, Rayciska (Fujiyama Haru)

- (11) *nee carehaw kii wa i-nuu-re kane*
some story do and INDEF.O-hear-CAUS FIN

‘Tell us a story, please.’ — Furukawa 1967: 106; West Sakhalin, Rayciska (Fujiyama Haru)

Furukawa lists *i-* as the first-plural object prefix; we treat it as the object form of the indefinite set (Chapter 13).

Hortatives use the indefinite-person suffix *-an* on the plural verb stem, with the invitation particles *anaa* (polite) or *roo* (familiar) (Furukawa 1967: 108) (Chapter 15):

- (12) *keh paye-an roo*
come.on go.PL-INDEF.S FIN

‘Come on, let’s go!’ — Furukawa 1967: 108; West Sakhalin, Rayciska (Fujiyama Haru)

19.6 The final-particle slot

Furukawa treats the sentence-final particle as an obligatory constituent of the clause, realised as zero (with its intonation) when no overt particle occurs (Furukawa 1967: 105). Her Rayciska inventory, the fullest we have for any Sakhalin variety, is worth setting out; the modal particles overlap with the system described in Chapter 15 and the discourse particles with Chapter 24.

Particle	Use
<i>kane</i>	polite request
<i>wa</i> (rising)	plain command
<i>wa</i> (falling)	question
<i>kah</i>	brusque command to intimates
<i>anaa</i>	polite invitation, ‘let’s’
<i>roo ~ roh</i>	strong invitation among equals
<i>nanko</i>	defiant refusal, ‘as if I would!’
<i>neya</i>	counterfactual wish, ‘if only’
<i>(kusu) iki</i>	confident assertion about the future
<i>naa</i>	confident warning or prediction
<i>wanaa</i>	‘there, just as expected’
<i>noo</i>	exclamation, admiration
<i>okaa</i>	wish, ‘I would like’
<i>aa</i>	emphatic question on a perfect form

Table 37: Final particles of Rayciska (after Furukawa 1967: 105–110)

Negative indefinite clauses — built with *ne(e) ... ka* plus a negative predicate — round out the picture of simple-clause types (Chapter 17):

- (13) *ne aynu ka isam*
 any person even not.exist

‘There is nobody at all.’ — Furukawa 1967: 107; West Sakhalin, Rayciska (Fujiyama Haru)

Chapter 20 · Nominalization

Nominalization with *pe/p*, *usi*, *hi*, and the possessive-suffix strategy; headless and lexical nominalizations.

Sakhalin Ainu turns clauses into noun phrases in two ways. The first it shares with the rest of Ainu: a nominalizing word such as *pe* ‘thing’ or *hi* ‘thing, time’ closes the clause. The second is its own innovation, and the centrepiece of this chapter: the possessive suffix of nouns (Chapter 7) attaches to verbs, creating a non-finite nominalized form that serves in adverbial clauses, under modal auxiliaries, as a clausal argument, in the evidential construction, and finally as an independent sentence form. Bugaeva analysed this *V*-possessive form and its history in the study this chapter follows throughout (Bugaeva 2016).

Because the chapter turns on the possessive form, its functional range is worth setting out at the start. The distribution is sharp: the nominalized verb appears in four constructions, and it is just as systematically excluded from three others (Bugaeva 2016: 105–106).

Construction	V-POSS?	Strategy used	Treated in
Clausal argument	yes	V-POSS	§20.2
Adverbial clause	yes	V-POSS + copula-based conjunction	§20.3
Evidential construction	yes	[clause] N-POSS <i>an</i> ($\pm$ V-POSS <i>nee</i>)	§20.4
Independent main clause	yes	V-POSS alone (insubordination)	§20.5
Relative clause	no	prenominal gap	Chapter 21
Headed complement clause	no	<i>pe</i> , <i>hi</i> , <i>kun-i</i>	Chapter 22
Lexical (participant) nominalization	no	<i>pe</i> , <i>hi</i> , <i>kur</i> , <i>usi</i>	§20.1

Table 38: Functional distribution of the possessive nominalization (after Bugaeva (2016: 112), Table 1)

The negative half of the table is as diagnostic as the positive. Noun modification never recruits the form: relative clauses keep the inherited prenominal gap strategy throughout the dialect area (Chapter 21). Lexical nominalization — making words for persons, things, and places — remains the preserve of the bound nouns of §20.1. And the complement clauses that a predicate selects with an overt nominalizing noun likewise resist it (Bugaeva 2016: 105). Where a possessive-nominalized clause does stand as the argument of a higher predicate, as in (3) below, it does so as a clausal nominalization in argument function — the construction that §22.5 examines from the predicate's side. The innovation thus spread along a single path, from possessed noun through adverbial and evidential scaffolding to the main clause, without ever displacing the older strategies from the territory they already held.

20.1 Nominalizing words

The bound nouns *pe* ~ *p* ‘thing, one’ and *hi* ~ *i* ‘thing, place, time’ head participant and event nominalizations. With *pe* the result ranges from fully lexicalized nouns to transparent headless relatives ‘the one that...’ (Chapter 21):

- (1) *etuhka ciwteh-pe ne yawa*
crow serve-NMLZ COP FIN

‘He is the emissary of the crow, you see!’ — Murasaki 2001: 198; Dal Corso 2025b: 49; West Sakhalin, Odasu (Asai Take)
ciwtehpe ‘servant, emissary’, a lexicalized *pe*-nominalization.

- (2) *sine cep ka isam-hi neampe*
one fish even not.exist-NMLZ TOP

‘as for the fact that there was not a single fish...’ — Pilsudski 1912: 196; Dal Corso 2025b: 53; East Sakhalin

Event nominalizations in *hi* as in (2) feed topicalization, complementation, and the conditional-irrealis form *kun(i) pe* (Chapter 22). The same bound nouns are fossilized inside core grammar: the West-coast negator *hanneh(ka)* contains *ham=ne-p* ‘NEG=COP-thing’ (§17.1), and the topic particle *neampe* is a frozen *ne-an-pe* ‘the thing which is’ (Chapter 24). The place noun *usi(ke)* ‘place’ supplies locational nominalizations, surviving in lexicalized place-words; productive examples in the Sakhalin record are scarce, and locational nominalization is normally carried by *hi* or by relativization on a locative noun (Chapter 21).

20.2 The possessive nominalization

Sakhalin Ainu extended the possessive form beyond the noun. Where Hokkaidō varieties possess only a subset of nouns, Sakhalin generalized the suffix to all nouns — *an=mosir-ih* ‘our country’ (Murasaki (1979: 5); (Bugaeva 2016: 105)) — and then carried the same morphology onto verbs, yielding a genuinely non-finite form *V-POSS* that has no counterpart elsewhere in Ainu (Bugaeva 2016: 104–105). The suffix copies the final vowel of the stem with epenthetic *h*, exactly as on nouns (§5.2). Murasaki had recorded these verb forms as a separate particle *hVV* of assertion and sympathetic questioning (Murasaki 1979: 72); Bugaeva identified them with the possessive paradigm and unified their uses (Bugaeva 2016: 112). We adopt her analysis, and gloss the form *POSS* throughout the grammar.

As a clausal argument, the nominalized verb stands where a noun phrase would:

- (3) *isa oh-ta ku=oman keray-kusu ku=araka-ha pirika*
doctor place-LOC 1SG=go.SG thanks-because 1SG=be.sick-POSS be.good

‘Because I went to the doctor, my sickness got better.’ — Murasaki 1979: 141; Bugaeva 2016: 108; West Sakhalin, Rayciska (Fujiyama Haru)
ku=araka-ha ‘my being sick’ is the subject of *pirika* ‘be good’.

20.3 Adverbial and modal hosts

Most occurrences of the possessive nominalization are governed by a small set of conjunctions and auxiliaries that contain the copula *ne(e)*: manner *neeno* ‘as, like’, adversative *neya* ‘but’, concessive *neyah* ‘although’, affirmative *neeko* ‘surely’, dubitative *nean* and *nee nankor* ‘probably’ (Bugaeva 2016: 106–107). Their Hokkaidō cognates, where they exist, govern finite verbs; the Sakhalin pattern *V-POSS* + *ne*-form is the innovation.

- (4) *ene e=ramu-hu nee-no pirika-no kii wa*
 like.this 2SG=think-POSS COP-ADV be.good-ADV do FIN

‘Do it well, the way you think.’ — Murasaki 1979: 139; Bugaeva 2016: 106; West Sakhalin, Rayciska (Fujiyama Haru)

- (5) *an=nukara-ha ne-ya nani hempah isam*
 INDEF.A=see-POSS COP-Q soon how.many not.exist

‘I saw it, but it was gone at once.’ — Murasaki 1979: 139; Bugaeva 2016: 106; West Sakhalin, Rayciska (Fujiyama Haru)

- (6) *pohke-no mokoro anah pirika-ha nee nankor*
 warm-ADV sleep if be.good-POSS COP probably

‘It will probably be good if he sleeps warm.’ — Murasaki 1979: 95; Bugaeva 2016: 107; West Sakhalin, Rayciska (Fujiyama Haru)

20.4 Nominalization in the evidential construction

The evidential sentences of Chapter 16 are themselves built on nominalization: the possessed evidential nouns *ruwehe*, *hawehe*, *sirihi*, *humih* head the clause, taking the scope proposition as their possessor and existential *an* as predicate, on the template [possessor clause]–N-POSS–*an*. In the basic pattern the possessor is a plain finite clause and no overt nominaliser appears on the scope verb at all — *re monimahpo an ruwehe an* ‘evidently three girls lived there’, literally ‘the trace of three girls being there existed’ (§16.3). Sakhalin diverges from Hokkaidō in also admitting a copula-linked subtype, in which the lexical verb is overtly nominalised with the possessive suffix and joined to the evidential noun by the copula *nee* — V-POSS *nee* N-POSS *an* (Bugaeva 2016: 108):

- (7) *tara aynu itah-no-ho nee sir-ih*
 that person speak-much-POSS COP look-POSS exist.SG

‘It looks like this person talks too much.’ — Murasaki 1979: 46; Bugaeva 2016: 108; West Sakhalin, Rayciska (Fujiyama Haru)

20.5 Main-clause nominalizations

The possessive nominalization also stands sentence-finally with no copula and full illocutionary force — an insubordination pattern. Bugaeva’s survey of conversation finds it in assertions, polar questions, content questions, and exclamations, and identifies the common thread as presupposition: the nominalized form presents its content as established or taken for granted, while finite forms carry new information (Bugaeva 2016: 109–112):

- (8) *kuani tani uwas ku=ek-ihii*
 1SG now just 1SG=come.SG-POSS

‘I have just come.’ — Murasaki 2013; Bugaeva 2016: 109; West Sakhalin, Rayciska (Fujiyama Haru)

- (9) *siriman poro-n-no eci=tarap-ihii*
 last.night many-EP-ADV 2PL=dream-POSS

‘Did you dream a lot last night?’ — Murasaki 2013; Bugaeva 2016: 109; West Sakhalin, Rayciska (Fujiyama Haru)

- (10) *hemanta kusu e=cis-ih*
 what reason 2SG=cry-POSS that boy

‘Why are you crying, boy?’ — Murasaki 2013; Bugaeva 2016: 109; West Sakhalin, Rayciska (Fujiyama Haru)

A content question with an afterthought vocative; on the postverbal noun phrase see §19.2.

Furukawa’s Rayciska data show the same form under her ‘past’ reading — *tara aynu omanihi* ‘that person went / did he go?’ — and the question particle *aa* selects precisely this nominalized form (Furukawa (1967: 109–110); §19.4). The perfect-like flavour follows from the presuppositional semantics: an event presented as settled fact.

That the sentence-final form really is a nominal, and not a finite verb that happens to resemble one, can be shown rather than assumed. The interrogative final particle *hetaneya* attaches only to nouns; yet it accepts the nominalized verb as its host, as in the rhetorical question *an-i-hi hetaneya* ‘why on earth is there...?’ of §16.6. Dal Corso reads this as the decisive diagnostic and analyses the sentence-final V-POSS (and V-POSS + copula)

patterns as true insubordination: formally subordinate clauses conventionalized as independent assertions of personal knowledge (Dal Corso 2018: 134–135).

The main-clause use is also where the two coasts part ways. The Rayciska record shows the full range of the table in the chapter introduction, insubordinated assertions and questions included. The East-coast texts deploy the nominalization freely in its embedded functions — as a clausal adjunct under causal *kusu*, e.g. *naha nuu-hu kusu* ‘because of his hearing that’, and as the object of a cognition verb, e.g. the ...-*pi-hi* nominalization embedded under *an-eramuskari* ‘I do not know’ (Pilsudski 1912) (Chapter 22) — but a bare sentence-final assertion or question of the West-coast type is not clearly attested in them. We read the asymmetry diachronically: the embedded uses belong to the older stages of the grammaticalization path traced in §20.6, the insubordinated main clause to the youngest, so the less innovative East retains the construction precisely where it arose first. A securely main-clause East-coast token would overturn this reading; we have not found one, and we offer the gap as a provisional dialectal finding rather than an accident of the record.

20.6 History

Bugaeva orders the Ainu nominalization strategies into a grammaticalization sequence: zero-nominalization of the bare verb is the oldest layer; the nominalizing-word strategy of Hokkaidō is the second; and the Sakhalin possessive nominalization, first in copula-based adverbial clauses, then in the evidential construction, and finally as a main-clause form, is the youngest (Bugaeva 2016: 113). All three layers coexist in Sakhalin Ainu, which thus preserves the history of the category in its synchrony. She raises the possibility that the non-finite verb form arose under the influence of the Tungusic languages of Sakhalin, with their fully developed non-finite subordination (Bugaeva 2016: 114); the contact setting is real (§1.4), and we consider the hypothesis plausible but not yet demonstrated — settling it would require dating the rise of the verb-possessive forms against the chronology of Ainu–Uilta bilingualism.

Chapter 21 · Relative clauses

Prenominal relativization, accessible positions, and the relation between relative clauses and noun incorporation.

Relativization in Sakhalin Ainu is simple in form and pervasive in use: the modifying clause stands before the head noun, the relativized argument is gapped, and there is no relativizer, no special verb form, and no agreement between clause and head. The verb of the relative clause keeps its finite shape, including person marking — the possessive nominalization of Chapter 20 never appears here (Bugaeva 2016: 98–99, 105–106). Furukawa’s phrase-structure rule for Rayciska places the modifying clause between numeral and head, inside the noun phrase template of §19.3 (Furukawa 1967: 102).

The gapped relative clause is the central member of a wider family of noun-modifying constructions. Sakhalin also builds gapless ones, in which the modifying clause relativizes no argument of its own but stands as possessor to a perception noun — ‘the trace / sight / sound of [clause]’ (Bugaeva 2017: 220–221); that type has been grammaticalized into the evidential system and is treated with it in §16.3. This chapter covers the gapped type.

21.1 The gap strategy

Any clause may modify a noun. The simplest case is a bare intransitive verb — the ‘adjectives’ of Chapter 6 (Word classes), as in Furukawa’s *kurasno seta* ‘black dog’, are exactly this construction with a one-word clause. With fuller clauses the pattern is unchanged:

- (1) *ehah-taa monimahpo utah hecire-hci wa sap-ahci*
 lily.root-dig girl PL dance-PL and come.down.PL-PL

‘The girls who went lily-root digging are coming down dancing.’ — Murasaki 2001: text 5; West Sakhalin, Odasu (Asai Take)

The modifier *ehah-taa* is itself a noun–verb compound (§14.6); the relativized argument is the S of ‘dig’.

The head noun fills its ordinary slot in the main clause, and the relative clause may in turn contain the person prefixes that its own arguments require. When the head corresponds to the object of the modifying clause, the clause-internal A marking stays in place:

- (2) *ci-reske* *moromahpo* *kam-ih* *ci-ee* *rusuy*
 1PL.EXCL-raise girl flesh-POSS 1PL.EXCL-eat DES

‘we want to eat the flesh of the girl we raised’ — Murasaki 2001: text 36; West Sakhalin, Odasu (Asai Take)
cireske moromahpo ‘the girl whom we raised’; the head is further possessed (*kamihi* ‘her flesh’).

- (3) *opista* *ci=nukara* *aynu* *uta*
 all 1PL.EXCL=see person PL

‘all the Ainu people whom we met’ — Tangiku & Ogihara 2001: Letter 3; East Sakhalin, Otasan (Sentoku Tarōji)

21.2 Accessible positions

Measured against the accessibility hierarchy, Ainu relativization reaches nearly the whole way down: subjects, objects, obliques, and inalienable possessors can all head a relative clause, and only the alienable possessor and the standard of comparison are excluded (Bugaeva 2015: 43). Within that wide reach the corpus distributes its heads very unevenly. S and O relativize freely and account for nearly all heads; (1) and (2)–(3) illustrate each, and existential relatives of the shape *cise ohta an N* ‘the N who was in the house’ are among the most frequent of all (§18.3). Heads corresponding to A are structurally unproblematic — the modifying clause simply retains its overt object — but they are rare in running text, which we attribute to discourse rather than grammar: A referents are typically topical protagonists, person-marked on the verb and not in need of restrictive modification. The pattern matches the cross-linguistic skew of relativization frequency along the accessibility hierarchy rather than any Sakhalin-specific constraint. The lower positions — obliques, possessors, and the standard of comparison — each recruit extra machinery, and the next three sections take them in turn.

21.3 Obliques: retaining the role marker

Below the core roles the bare gap stops working: verb agreement registers arguments, not adjuncts, so an oblique gap would leave its role unrecoverable. The Ainu answer, family-wide, is retention — the postposition that would have marked the constituent stays behind inside the relative clause, orphaned of its complement, and signals the role of the head. In Hokkaidō the device is textbook material; Satō’s paradigm pair shows the instrumental *ani* left at the clause edge once its complement *makiri* ‘knife’ is promoted to head (Satō 2008: 263–264):

- (4) *makiri* *ani* *cep* *ku=tuye*
 knife INS fish 1SG.A=cut

‘I cut fish with a knife.’ — Satō 2008: 263–264; Hokkaidō

- (5) *ani* *cep* *ku=tuye* *makiri*
 INS fish 1SG.A=cut knife

‘the knife I cut fish with’ — Satō 2008: 263–264; Hokkaidō

The orphaned instrumental is the only signal of the head’s role inside the clause; no resumptive element appears.

Not every marker can be orphaned. Bugaeva delimits the strandable set as the instrumental *ani*, the comitative *tura*, the dative *eun*, and the terminative *pakno*; the plain locative *ta* and allative *un* cannot strand, and a spatial adjunct relativizes instead through a relational noun such as *oro* ‘place of’, which remains in the clause carrying the locative marker (Bugaeva 2015: 48, 50). The Sakhalin counterparts of the strandable set are *ani* ‘with, by means of’, *tura* ‘together with’, and the terminative *pahno* ‘until, as far as’ (§11.1), the dative work of Hokkaidō *eun* falling to *onne*.

The one direct Sakhalin token of an oblique relative obeys the constraint exactly. In (6) a goal of motion has been relativized, and the clause does not strand bare *ta*: it retains *or-ta* ‘at the place (of)’ — the relational-noun bridge — which Dal Corso analyses as the cue that keeps the role of the gapped constituent recoverable (Dal Corso 2025b: 53):

- (6) *or-ta ariki-an-kun Hokkayto mosiri*
place-LOC come.PL-INDEF.S-COND Hokkaidō land

‘Hokkaidō, to which we would (never have thought of) coming’ — Murasaki 1976: 5; Dal Corso 2025b: 53; West Sakhalin, Rayciska

The retained locative *orta* marks the relativized constituent as a goal; *-kun* marks the clause as irrealis (§22.2).

Spatial relativization also has a route that needs no retained marker at all: a place noun as head. In Sentoku's letters the locus argument of *an* ‘exist’ ‘be (at)’ is gapped under the head *toko* ‘place’, whose own semantics supply the spatial role; the headed phrase then behaves like any noun, here as complement of the terminative (the same token quoted in §11.4):

- (7) *israh kara wa cokay e=an toko pak no paye=an kusu iki*
message make and 1PL 2SG=exist.SG place until ADV go.PL-INDEF.S PURP do

‘I will send word, and we will travel as far as the place where you are.’ — Tangiku & Ogihara 2001: Letter 1; East Sakhalin, Otasan (Sentoku Tarōji)

e=an toko ‘the place where you are’: the locative argument of *an* is gapped, and the place-noun head makes a stranded *ta* unnecessary.

Retention competes, finally, with the applicatives of §14.2. An oblique promoted to object by *e-*, *ko-*, or *o-* relativizes by the plain gap, like any other object; for Hokkaidō, Bugaeva shows the applicative route taken wherever the lexicon provides it, with retention as the fallback for roles no applicative covers (Bugaeva 2010: 760). The Sakhalin texts give us no secure relative clause headed by an applied object, but the feeding relation costs nothing: once *e-tuhse-ka* ‘kick with’ has made the instrument its object, nothing distinguishes that object from any other O under relativization.

21.4 Possessors

An inalienable possessor relativizes with nominal rather than adpositional tracking material: the possessum stays inside the clause in its possessed form, and the retained possessive suffix is what identifies the head as possessor (Bugaeva 2015: 49). Sakhalin is structurally well prepared for this, having generalized the possessed form across nearly the whole noun lexicon (§7.1). The texts show the pattern in phrasings built from clauses like *maci-hi isan* ‘his wife was away’ (§17.4): in *macihi isan horokewpo*-type phrases, the possessive morphology on *maci-hi* is all that marks *horokewpo* ‘young man’ as the one whose wife is meant. We analyse this as the regular gap strategy applied to a clause whose subject is a possessed noun — no further machinery is needed. The alienable possessor, by contrast, is one of the two genuinely inaccessible positions (Bugaeva 2015: 43): ‘the man whose dog...’ has no direct rendering, and the texts paraphrase through the appositive possession strategy of §7.4.

21.5 The standard of comparison

The other inaccessible position is served by a conventional work-around rather than a relative clause proper. The standard of a comparison (‘than X’) cannot itself be relativized; what relativizes instead is a degree adjunct in the terminative *pakno* ‘up to, as far as’, stranded like the other members of the strandable set, with the negative existential *isam* ‘not exist’ as the clause's predicate: ‘a Y [up to whom no Y exists]’, that is, ‘a Y without equal’ — a superlative (Bugaeva 2015: 52–53). Both the idiom and the equative comparison underlying it are on record from Hokkaidō (Saru):

- (8) *a=ona sorekusu pak-no nispa isam nispa ne*
INDEF.A=father truly until-ADV rich.man not.exist rich.man COP

‘My father was a rich man truly without equal.’ (lit. ‘my father was a [no rich man (reaching) up to him] rich man’) — Bugaeva 2015: 53; Hokkaidō (Saru)

The degree adjunct is relativized and *pakno* stranded before the negative existential; the matrix clause is equative.

- (9) *ne tures-i pak pirka menoko isam*
that younger.sister-POSS until be.good woman not.exist

‘There was no woman as fine as his younger sister.’ — Bugaeva 2015: 53; Hokkaidō (Saru)

The equative base of the superlative frame: standard + *pak* + parameter, with existence negated.

No Sakhalin token of the complete frame survives, but each of its ingredients is in place: the terminative *pahno* ‘until, as far as’ is frequent (§11.4) and takes a gap-headed relative as its complement in (7); *isam* ‘not exist’ is the everyday negative existential (§17.4); and degree adjuncts belong to the strandable set. We accordingly read the absence as a gap in the record rather than in the grammar — though, as elsewhere, confirmation would take a token of the shape [X *pahno* Y *isam*] Y *ne*, which the surviving texts do not provide.

21.6 Headless relatives

With the bound nouns *pe* ~ *p* ‘one, thing’ and *hi* ‘thing, time, place’ as heads, relativization yields the participant and event nominalizations of §20.1 — ‘the one that...’, ‘the fact that...’. These heads also combine with the possessive suffix, as in *an=oyra-pe-he* ‘the things we had forgotten’ (Murasaki (1979: 51); (Bugueva 2016: 106)). The chain noun > bound noun > nominalizer is the same grammaticalization that produced the negator *hanneh(ka)* and the topic particle *neampe* (§17.1; Chapter 24).

21.7 Relative clauses in the noun phrase

Within the noun phrase, the relative clause follows demonstratives and numerals and immediately precedes the head (§19.3); demonstrative plus relative is common in narrative for re-identifying referents (*neya cise ohta an monimahpo*-type phrases; §9.3). The corpus gives no evidence of a restrictive–non-restrictive contrast in form; both readings are carried by the same construction and resolved in context. Hokkaidō Ainu allows relativization to interact with noun incorporation in ways that have attracted theoretical attention (Dal Corso 2020); the Sakhalin corpus offers parallel ingredients — compare the incorporated modifier in (1) — but no studied cases, and we leave the interaction as an open question: settling it would require elicitation or a larger text base than survives.

Chapter 22 · Complementation

Complement clauses with *kuni*, *hi/humi/hawe/ruwe* and bare complements; complement-taking predicates.

Sakhalin Ainu has no complementizer of the English *that* type. Clauses become arguments of predicates in four ways: as bare, zero-nominalized complements of modal and aspectual predicates; as irrealis complements built with *kun(i)* and the bound nouns *pe* and *hi*; as direct quotations closed by the quotative *nah* with a verb of saying; and — in the distinctively Sakhalin pattern — as possessive-nominalized clauses, which serve perception predicates as objects and evaluative predicates as subjects, shading into the evidential constructions of Chapter 16. Throughout, complementation leans on the nominalization machinery of Chapter 20. In the terms of Noonan’s typology of complementation, the language divides the labour between a paratactic direct-quote type, a reduced subject-sharing type, and nominalized complements, with no finite indicative complement clause at all (Noonan 2007).

Strategy	Shape	Selecting predicates	Section
Zero-nominalization	bare [V]; person marking migrates to the matrix	<i>rusuy</i> ‘want’, <i>easkay</i> ‘be able’, <i>koyaykus</i> ~ <i>eaykah</i> ‘cannot’, <i>hemaka</i> ‘finish’, <i>etunne</i> ‘not want’, <i>erameskari</i> ‘not know’	§22.1
Irrealis <i>kun(i)</i>	[V <i>kun-i</i>], [V <i>kun pe</i>], purposive <i>kuni no</i>	<i>ramu</i> ‘think, intend’; future and obligative periphrases	§22.2
Quotation	[direct clause] <i>nah</i> + framing verb	<i>yee</i> ‘say’, <i>kii</i> ‘go (of speech, sounds)’, <i>ramu</i> ‘think’	§22.3
Possessive nominalization	V-POSS (± <i>neampe</i> , ± <i>nee</i>)	perception verbs as object; evaluative <i>pirka</i> ‘be good’, <i>sunke</i> ‘be untrue’ as subject	§§22.4–22.5

Table 39: Complementation strategies and the predicates that select them

The chapter takes the strategies in turn and closes with the selectional system that unites them (§22.6).

22.1 Bare complements of modal and aspectual predicates

The desiderative *rusuy*, potential *easkay* and its negatives *koyaykus* and *eaykah*, and completive *hemaka* take a bare verb as their complement — a zero-nominalization, the oldest layer of the system (Bugueva 2016: 99, 113) (Chapter 15 for their semantics). Furukawa's Rayciska data show the diagnostic property: the person marking of the whole sentence migrates to the modal predicate, leaving the notional verb bare — her transformational rule reorders *ku-oman koyaykus* into the usual

- (1) *oman ku-koyaykus*
go.SG 1SG-cannot

‘I cannot go.’ — Furukawa 1967: 113; West Sakhalin, Rayciska (Fujiyama Haru)

Beside less common *ku-oman koyaykus* and doubly marked *ku-oman ku-koyaykus* (Furukawa 1967: 112 fn. 5).

- (2) *ampene suukawka ku-koyaykus*
at.all sew 1SG-cannot

‘I cannot sew at all.’ — Furukawa 1967: 106; West Sakhalin, Rayciska (Fujiyama Haru)

koyaykus ‘cannot (because of circumstances)’ contrasts with *eaykah* ‘cannot (by inability)’ (Furukawa 1967: 106).

The complement may be flagged with the additive particle *ka* ‘even’, and with the lexically negative predicates of §17.4 this yields fixed constructions: *V ka erameskari* ‘have never V-ed’ (experiential), *V ka etunne* ‘not want to V’, *V ka koyaykus* ‘be unable to V’ (Dal Corso 2025b: 40–41). The negated light-verb construction of §17.2 — nominalized verb + *ka* + *hanki* — has exactly this complement syntax, with ‘do’ as the matrix predicate.

22.2 Irrealis complements with *kun(i)*

Unrealized, intended, or obligatory events are embedded with *kun(i)*, which we gloss COND following its conditional-irrealis value; it combines with the nominalizers as *kun-i* and *kun pe* (§20.1). Under verbs of thinking and intending the *kun-i* form supplies the complement:

- (3) *ene an-ki-kun-i an-ramu*
like.this INDEF.A-do-COND-NMLZ INDEF.A-think

‘I intended to do so.’ — Pilsudski 1912: 210; Dal Corso 2025b: 53; East Sakhalin

With the adverbializer *no*, the same form makes purposive complements and adjuncts, ‘so that, in order to’; the oldest West-coast texts show it under negation:

- (4) *ukoramuwen ham an-ki kuni no*
quarrel NEG INDEF.A-do COND ADV

‘so that we do not quarrel (any more)...’ — Dobrotvorsky 1875; Sakaguchi 2021; West Sakhalin, West coast, central (recorded 1867–1872)

The same texts show the assertive future complex *nankoro pe ne* ‘it will (probably) be the case that...’, where the dubitative, the bound noun *pe*, and the copula stack over the clause — complementation grammaticalized into modality (Chapter 15):

- (5) *irutaspano ukoramupirika an-ki ciki pirika nankoro pe ne*
mutually reconcile INDEF.A-do if be.good probably NMLZ COP

‘If we make peace with each other, it will be a good thing.’ — Dobrotvorsky 1875; Sakaguchi 2021; West Sakhalin, West coast, central (recorded 1867–1872)

22.3 Quotation and reported speech

Speech and thought are reported as direct quotation: the quoted clause stands first, the quotative particle *nah* closes it, and a verb of saying or doing follows — *nah yee* ‘say so’, *nah kii* ‘go like this’ (of speech and sounds), *nah ramu* ‘think so’. Inside quotations the deixis is that of the original speaker, and in folktales the quotative frame carries the logophoric use of the indefinite person (§13.2; §16.7):

- (6) *yee wa an-nuu nah yehci yahka ampene mokoro manu*
 say and INDEF.A-hear QUOT say.PL although utterly sleep REP

“Tell us, and we will listen,” they said — but he just slept on.’ — Murasaki 2001: text 5; West Sakhalin, Odasu (Asai Take)

yehci < *yee* + *-hci* with regular vowel shortening (§5.4); the quotation is a command with its own final particle omitted in the retelling.

That logophoric convention deserves spelling out, because it is the closest the language comes to a dedicated reported-speech grammar. When a character reports speech in which he or she figures, the quoted speaker is encoded with the indefinite (fourth-person) prefixes rather than with the first person: the “I” inside the quote is grammatically held apart from the “I” of the narrator. Chiri noticed the effect and remarked that the Sakhalin indefinite inside quotations comes close to a third-person value — so close that he was inclined to call it a third-person affix outright (Chiri 1942: 764) — citing a cry from Pilsudski’s East-coast texts in which the quoted speaker’s ‘me’ is the indefinite object prefix:

- (7) *oha-cisuye i=nospa ike kira yan nah ye manu*
 empty-house.demon INDEF.O=chase and.so flee IMP.PL QUOT say REP

“The demon of the empty house is chasing me — run!” he said.’ — Pilsudski 1912: 100; Chiri 1942: 764; East Sakhalin

The quoted speaker refers to himself with the indefinite object prefix *i* = ‘me’, where first-person *en* = might be expected; the quotative frame licenses the shift.

Sakhalin Ainu thus has no logophoric pronoun series of the West African type — no form reserved for the reported speaker and nothing else. What it has is the independent fourth-person category recruited for the logophoric function, disambiguating the speech-internal from the speech-external participant without new morphology. The recruitment is the diachronically expected one: indefinite and fourth-person categories are a recognized source of logophoric marking cross-linguistically, and the Sakhalin texts show the source construction still transparent, quasi-logophoric rather than grammaticalized (§13.2).

Indirect speech as a distinct construction is not attested: even reports embedded under *nah ramu* ‘think’ keep direct-speech deixis. The reportative *manu* that ends narrative sentences is the same strategy scaled up to the whole text — the tale is one long quotation (Chapter 16).

22.4 Perception and knowledge complements

Perception verbs take a nominalized clause as object. The commonest frame in the narrative corpus is the possessive nominalization plus topic particle, *V*-POSS *neampe*, literally ‘(as for) my seeing — ...’, which serves at once as a perception report and a temporal pivot ‘when I looked’:

- (8) *an-nukara-ha neampe suy hekimoh makanu wa isam manu*
 INDEF.A-see-POSS TOP again uphill go.up and not.exist REP

‘When I looked — it had gone off uphill again.’ — Murasaki 2001: text 5; West Sakhalin, Odasu (Asai Take)

We analyse the construction as a nominalized complement that has been topicalized out of its matrix clause: the perceived content is then asserted as a fresh main clause, exactly parallel to the presuppositional main-clause nominalizations of §20.5. Where the perceived content is expressed with the evidential nouns — ‘hear the voice of X V-ing’, ‘see the sight of...’ — the complement structure and the evidential construction coincide (§16.4; (Bugaeva 2016: 101–102)). Knowledge predicates pattern with the modals: *wante* ‘know’ and its negative *erameskari* take bare or *ka*-marked zero-nominalized complements (§22.1), as in the experiential construction *V ka erameskari* ‘have no experience of V-ing’ (Dal Corso 2025b: 40–41). The East-coast texts show that the cognition class also reaches the possessive nominalization: Pilsudski’s *yay-ikoramopi-hi kayki an-eramuskar* ‘I do not even know what (she) thinks of me’ embeds a *V*-POSS form as the object of the negative knowledge verb (Pilsudski 1912), so the strategy is not confined to the innovating West coast in its embedded uses (§20.5).

22.5 Evaluative predicates and clausal subjects

A clause can also be the *subject* of a higher predicate. The selecting class is the one-place evaluative and veridicality predicates — *pirka* ‘be good’, *wen* ‘be bad’, *sunke* ‘be untrue, lie’ — which predicate something of a state of affairs rather than of a participant. On the West coast that state of affairs is packaged as the

possessive nominalization: in *ku=araka-ha pirika* ‘my being sick got better’, the clause *ku=araka-ha* occupies the subject slot of *pirika* (the example is given in full in §20.2), and in *ku=ye-he sunke* ‘what I said is untrue’ the nominalized clause is what the veridicality predicate judges false (Chapter 7; Murasaki (1979: 95, 141); Bugaeva (2016: 108)).

The construction is the structural opposite of the modal class of §22.1. There the complement is stripped bare and the person marking migrates onto the matrix predicate; here the complement keeps its own person marking and carries full nominalizing morphology, while the matrix contributes nothing but a plain one-place predicate. The two classes show no attested overlap: no evaluative predicate is attested with a bare-verb complement, and no desiderative or potential with a V-POSS one. The attested evaluative class is admittedly small — only *pirika* and *sunke* are secure in the corpus as matrices of a nominalized subject clause. Its semantics, however, defines a natural class, and we predict that its remaining members, most obviously the antonym *wen* ‘be bad’, select the same V-POSS subject clause; a larger conversational sample would test the prediction.

22.6 The selectional system

Set side by side, the strategies of this chapter are not an arbitrary list: which one a matrix predicate selects follows from how tightly its meaning binds it to the complement event.

Predicate class	Predicates	Complement shape	Attested frame
Desiderative, modal, ability	<i>rusuy, easkay, koyaykus</i>	bare [V]	<i>oman ku-koyaykus</i> (§22.1)
Phasal	<i>hemaka</i>	bare [V]	<i>ray hemaka</i> (Chapter 15)
Evaluative, veridicality	<i>pirika, wen, sunke</i>	V-POSS subject clause	<i>ku=araka-ha pirika</i> (§22.5)
Evidential and modal noun heads	<i>ruwehe an, nee nankor</i>	V-POSS + <i>nee</i>	<i>pirika-ha nee nankor</i> (§20.3; Chapter 16)
Utterance	<i>yee, kii</i>	direct clause + <i>nah</i>	<i>... nah yee</i> (§22.3)
Cognition, knowledge	<i>ramu, wante, erameskari</i>	<i>nah</i> -clause, <i>kun-i</i> , or nominalization	<i>... kun-i an-ramu</i> (§22.2)

Table 40: Complement shape by matrix predicate class

The gradient is the iconic one of Givón's binding hierarchy (Givón 1980): the tighter the semantic bond between matrix and complement, the less nominalizing morphology the complement carries. The subject-sharing desideratives and modals, whose two events are in effect one, take a stripped verb welded into a single complex predicate; the evaluatives, which assert something about an independent but presupposed event, take a fully nominalized clause; and the utterance verbs, whose complement is an autonomous speech act, take an unreduced direct clause merely flagged by *nah*. The cognition verbs straddle the loose end of the scale, sharing the quotation frame with *yee* and the nominalizations with the evaluatives — as expected of predicates whose semantics ranges from quoted inner speech to settled knowledge. Within Noonan's typology, Sakhalin Ainu thus runs its entire complementation on sentence-like (quoted) and reduced or nominalized complement types (Noonan 2007); the finite indicative complement that English builds with *that* simply has no Sakhalin counterpart. The chief residual uncertainty is the size of the evaluative class (§22.5).

Chapter 23 · Clause linking

Conjunctive particles *teh*, *wa*, *kusu*, *ayne*, cooccurrence with *manu*, and clause chaining in narrative.

Sakhalin Ainu links clauses with a set of clause-final particles — *teh* and *wa* ‘and’, *ayne* ‘while, until eventually’, *kusu* ‘because, in order to’, *yahka* ‘although’, *anah* and *cikin* ‘if, when’, and others — and with a parallel set of sentence-initial connectives that carry narrative from sentence to sentence. Furukawa counted some forty clause-final linkers in Rayciska alone (Furukawa 1967: 110); their status between coordination and subordination has never been settled for the language (Dal Corso 2025b: 40, 49). We treat the system as clause

chaining: the linked verb keeps its person marking, takes no independent final particle or evidential, and shares illocutionary force with the clause that closes the chain — a cosubordination profile rather than either pole. Within the chain, the general ‘and’ linkers sort themselves along a single semantic scale, the temporal gap between the linked events (§23.1–23.2); the final section (§23.7) then reads Furukawa's forty against Chiri's catalogue and the corpus. The one linker that is formally bound to dependency is negative: caritive *hamo* ‘without doing’ marks its clause as non-finite all by itself (§17.6).

23.1 Sequential linkage: *wa*, *teh*, and the temporal gap

The two general ‘and’ linkers divide the territory by event structure, and Furukawa's elicitation captured the line between them. *wa* joins actions with no gap, so that the two fuse into a single event; *teh* sequences distinct actions with a perceptible interval between them; the narrative-succession linkers of §23.2 then push the second event into a new episode altogether. Her consultant accepted both of the following but rejected **ku-koro wa ku-eh*, where buying and coming are separate events (Furukawa 1967: 110–111):

- (1) *ku-koro-wa eci-konte*
1SG-buy-and 1SG>2SG-give

‘I bought it and gave it to you.’ — Furukawa 1967: 110; West Sakhalin, Rayciska (Fujiyama Haru)
On the portmanteau *eci*- ‘I...you’ see §13.4.

- (2) *ku-koro-teh ku-eh*
1SG-buy-and 1SG-come.SG

‘I bought it and came (here).’ — Furukawa 1967: 111; West Sakhalin, Rayciska (Fujiyama Haru)

Both linkers feed grammaticalization: *teh an* is the stative periphrasis and *wa isam* the conclusive (Chapter 15; §17.4). In Asai Take's narration *teh* is the default sequencer, often several to a sentence; the Sentoku letters use *wa* more freely. In *nukara konte wa* ‘please read it’ (Letter 1), though, the *wa* is the request-softening final particle that follows the benefactive, not a sequencer (Chapter 26).

The interval that *teh* asserts can be made explicit by a following *orowa* ‘after that’ — transparently the possessed relational noun *oro* ‘place (of it)’ plus the ablative *wa* ‘from’ (Chapter 11), so literally ‘from there’. The reinforced sequence *teh orowa* ‘and after that’ keeps the chain open for further links, here a purpose clause in *kusu* before the single finite close:

- (3) *e=niina teh orowa e=i-wooneka kusu teeta e=san-i-hi ne=an ike=aa*
2SG=gather.wood and.then after.that 2SG=INDEF.O- PURP here 2SG=go.down.SG- COP=PFV CONJ=FIN
check.on EP-PERS

‘You gathered firewood and then came down here to check on us, didn't you?’ — Dal Corso 2018: 210; West Sakhalin

teh orowa = explicit delayed sequence; *kusu* adds a purpose clause; the whole chain closes on one nominalized predicate with personal-knowledge *-hi* (Chapter 16).

23.2 Narrative succession: (h)ike and (h)ine

At the wide end of the temporal-gap scale stand the narrative-succession linkers. *ike* ‘and (so), when’ — historically the nominalizer *-(h)ike* ‘the one that...’ reanalysed as a linker — is the Sakhalin workhorse: Chiri notes the *h*-doublet and states that Sakhalin *(h)ike* serves ‘in the same sense as *wa*, *hine*, *tek*’, in which use it is best regarded as a conjunctive particle outright (Chiri 1942: §51, §61). In narrative it typically chains a backgrounded presentational clause onto the event line:

- (4) *re aynu okay-ahci ike ceh-ki-hci ko e-hci*
three person exist.PL-PL and fish-catch-PL whenever eat-PL

‘Three men were living (there), and whenever they caught fish they would eat them.’ — Kitahara 2016; West Sakhalin

ike chains the presentational existential to the main event line; on habitual-correlative *ko(h)* see Chapter 15. Segmentation ours.

Its partner *hine ~ ine*, which Chiri derives from *a-ine*, is the great narrative converb of Hokkaidō but thin in the Sakhalin texts, where *ike* — and its reportative fusion *manuyke* (§23.6) — does the corresponding work; *hine* survives most clearly in the stative periphrasis *hine an* that Chiri lists beside *teh an* (Chapter 15). The

comparative picture runs through *teh* itself: in Saru the cognate *tek* narrowed into an auxiliary ‘do for a moment’, leaving sequence to *wa* and *hine*, while Sakhalin *teh* keeps the older linking value (Bugaeva 2022). Chiri’s catalogue further records a counter-expectational *a(h)i ~ a(h)ike* ‘...when (unexpectedly), ...but’ built on the same formative (Chiri 1942: §39); the corpus does not let us draw its line against plain *ike*, a point taken up in §23.7.

23.3 Temporal and durative linkers

yayne marks a prolonged event overtaken by what follows — ‘while doing, until eventually’:

- (5) *ku-henoye yayne ku-mokoro hemaka*
1SG-doze while.eventually 1SG-sleep finish

‘While I dozed, I eventually fell fast asleep.’ — Furukawa 1967: 111; West Sakhalin, Rayciska (Fujiyama Haru)

The same family supplies the sentence-initial (*an*)*ayne* ‘in time, meanwhile’ of §23.6. *kane* links a simultaneous backgrounded action, ‘while’ — *kuwa koro kane ahkas* ‘walks holding a stick’ (§19.2) — beside its final-particle use in requests (§19.6). Posterior ‘after’ is built on the nominal *okaaketa* ‘in the aftermath of’, as in *macihi isan teh okaaketa* ‘after his wife had gone out’ (§17.4); anteriority needs no dedicated marker beyond *teh*.

23.4 Causal and purposive linkage

kusu covers cause and purpose, the reading resolved by the relation between the clauses; with the copula it builds the intentional future *kusu ne* and with *iki* the assertive future (Chapter 15). Nominal cause is carried by *renkayne* ‘owing to’, which also governs clauses, as in Sentoku’s *tani nuca isan renkayne* ‘now that the Russians are gone’ (§17.4). Negative purpose uses *kuni no* (§22.2). A single sentence of the Dobrotvorskyy-era texts stacks causal *ne kusu* with a concessive:

- (6) *sine utahta ne kusu porono ham e-ye yahka pirika nankoro pe ne*
one family COP because much NEG 2SG-say although be.good probably NMLZ COP

‘Since we are of one family, even if you do not say much, it will be well.’ — Dobrotvorskyy 1875; Sakaguchi 2021; West Sakhalin, West coast, central (recorded 1867–1872)

23.5 Conditionals and concessives

The conditional linkers are *anah(ka)* and *cik(in) ~ ciki*; both follow the bare protasis verb. *anah* shades from ‘if’ into topic marking, a polysemy taken up in Chapter 24:

- (7) *siripirika anah ku-oman kusu neya*
weather.be.good if 1SG-go.SG PURP FIN

‘If the weather were good, I would go.’ — Furukawa 1967: 108; West Sakhalin, Rayciska (Fujiyama Haru)
kusu plus the final particle *neya* expresses an unrealizable wish (Furukawa 1967: 108).

- (8) *sonko an=omante cikin nukara konte wa*
message INDEF.A=send if see give FIN

‘I send you word — when it reaches you, please read it.’ — Tangiku & Ogihara 2001: Letter 1; East Sakhalin, Otasan (Sentoku Tarōji)
konte after a verb is benefactive ‘do for someone’; *cikin* also appears in the contact phrase *somo ne ciki* ‘if not’ (§17.7).

The East-coast texts add the heavier protasis marker *kusu neyke*, which interacts with negation in the *hannah kusu* conditional construction (§17.5; (Dal Corso 2025b: 55–56)). The concessive is *yahka* ‘although, even if’, as in (6) and Furukawa’s *ekasre seeseh yahka wen* ‘even too much warmth is bad’ (Furukawa 1967: 111).

23.6 Narrative chaining

Connected narrative organizes these links into a recognizable rhythm. Clauses chain with *teh* and *ike* ‘and, when’ up to a finite verb capped by the reportative (Chapter 16); the next sentence then re-opens by repeating the last verb of the previous one — tail-head linkage, ubiquitous in Asai Take’s telling:

- (9) *otakaa-ta san manu. otakaa-ta san ike*
 beach-LOC go.down.SG REP beach-LOC go.down.SG and

‘He went down to the beach. Going down to the beach, (he swam...)’ — Murasaki 2001: text 5; West Sakhalin, Odasu (Asai Take)

The fusion of the reportative with *ike* yields the narrative linker *manuyke* ‘..., it is said, and’, which keeps a chain running across what is formally a sentence boundary:

- (10) *reske-hci yayne poro-hci manuyke*
 raise-PL while.eventually be.big-PL REP.and

‘they raised them and they eventually grew up, and then...’ — Murasaki 2001: text 19; West Sakhalin, Odasu (Asai Take)

Above the clause, a closed set of sentence-initial connectives — built from the copula, the existential, and *orowa* ‘from there, then’ — manages the larger discourse: *neeteh* and *nahteh* ‘and then’, *orowa(no)* ‘after that’, *anayne* ‘meanwhile, in time’, *neyke* ‘and so’. They open the majority of sentences in the Asai Take corpus, often doubled (*neeteh orowa tani...*); this inventory and its layering are our generalisation from the corpus. The same forms carry the topic-resumption work described in Chapter 24.

23.7 Furukawa's forty: reconstructing the inventory

Furukawa reports having identified some forty conjunctive forms in Rayciska but published worked examples of only four — *wa*, *teh*, *yayne*, *yahka* (Furukawa 1967: 108–111). The rest are not lost to description. Chiri's catalogue of conjunctive particles — his third particle class, drawn up pan-dialectally with the Sakhalin uses flagged (Chiri 1942: §120–121) — contains most of them, and reading that catalogue against the West- and East-coast texts recovers the bulk of the inventory. The table sorts the result by evidential footing: which linkers Furukawa exemplified, which the corpus confirms independently, and which still rest on Chiri's authority alone.

Form	Value	Footing
<i>wa</i>	‘and’; no temporal gap (§23.1)	published by Furukawa (Furukawa 1967: 110); pervasive in the corpus
<i>teh</i>	‘and then’; perceptible interval (§23.1)	published by Furukawa (Furukawa 1967: 111); Asai Take's default sequencer
<i>(h)ike</i>	‘and (so), when’; narrative succession (§23.2)	(Chiri 1942: §51, §61); corpus-confirmed, including tail-head chains and <i>manuyke</i> (§23.6)
<i>(h)ine</i>	narrative ‘and’	(Chiri 1942); thin in the Sakhalin texts, clearest in stative <i>hine an</i> (§23.2)
<i>a(h)i ~ a(h)ike</i>	‘...when (unexpectedly), ...but’	(Chiri 1942: §39); not separable from plain <i>ike</i> in the corpus
<i>yayne</i>	‘while, until eventually’ (§23.3)	published by Furukawa (Furukawa 1967: 111); corpus-confirmed
<i>(an)ayne</i>	‘in time, meanwhile’	corpus-confirmed; sentence-initial in Asai Take (§23.6; (Murasaki 2001))
<i>kane ~ kanne</i>	simultaneous ‘while, in the state of’	(Chiri 1942: §54); West-coast <i>kane</i> in chaining and requests (§23.3)
<i>koh (kor ~ ko)</i>	‘whenever’; habitual-correlative	(Chiri 1942: §54, §56, §60); corpus-confirmed (Chapter 15)
<i>orowa(no) ~ ora(no)</i>	‘and then, after that’	transparent <i>oro</i> ‘place’ + <i>wa</i> ‘from’; pervasive sentence-initially (§23.6; (Chiri 1942: §70–71))
<i>rapok(ta)</i>	‘while, in the interval’	Chiri's catalogue only (Chiri 1942: §120–121)
<i>kusu</i>	‘because; in order to’ (§23.4)	published by Furukawa (Furukawa 1967: 108); pervasive
<i>anah(ka)</i>	‘if, when’, shading into topic (§23.5)	published by Furukawa (Furukawa 1967: 108); pervasive
<i>ciki ~ cikin</i>	‘if, when’ (§23.5)	(Chiri 1942: §50); corpus-confirmed (Sentoku letters, (Tangiku & Ogihara 2001))
<i>yahka(yki)</i>	‘although, even if’ (§23.5)	published by Furukawa (Furukawa 1967: 111); corpus-confirmed
<i>korka</i>	‘but’	(Chiri 1942: §45); modern Sakhalin prose ((Kitahara 2016))
<i>pakno</i>	‘until, as far as’	corpus-confirmed as terminative postposition (§21.5); clause-linking use on Chiri's authority
<i>kunak ~ kuni</i>	irrealis (expectative) complementizer	(Chiri 1942: §64); see §22.2

Table 41: The Sakhalin clause-linking inventory: Furukawa's published four, the corpus, and Chiri's catalogue. On this reading Furukawa's ‘forty’ are no mystery: essentially the union of the linkers of this chapter with Chiri's catalogue, swelled by transparent reinforcements (*teh orowa*, *nahteh orowa*) and the copula-built connectives of §23.6. What remains genuinely open is the fine discrimination among near-synonyms — *a(h)i* against *a(h)ike*, *kane* against *koh* — in West-coast speech, which only elicitation could have settled. Two members of the catalogue belong to complementation rather than linking: the irrealis *kunak ~ kuni* (§22.2), and the quotative frame in *nah*, inside which a quoted speaker refers to him- or herself with the indefinite *an* rather than a first-person form — the quoted-speaker strategy described with quotation in §22.3.

Chapter 24 · Information structure

Topic marking with *anah(ka)*, focus particles, word-order variation, and discourse particles.

Grammatical relations in Sakhalin Ainu are carried by the verb (Chapter 13); noun phrases bear no case marking, and the postpositions of Chapter 11 mark only obliques. What a sentence is *about*, and which part of it is news, is therefore signalled by other means: a small set of particles placed after the phrase they scope over, departures from the verb-final default order of Chapter 19, and the final particles of dialogue. Chiri's classic description sorts the relevant particles into an adverbial class (*ka, kayki, patek...*), a Japanese-style "binding" class that includes the topic markers, and an interjectional class that may close the sentence (Chiri 1942: §124–§129). The data below come from Chiri's survey, Furukawa's Rayciska syntax (Furukawa 1967), Murasaki's conversation course (Murasaki 2025), Piłsudski's east-coast texts (Piłsudski (1912), with the analyses of Dal Corso (2018) and Dal Corso (2024)), the Sentoku Tarōji (千徳太郎治) letters (Tangiku & Ogihara 2001), and our own counts over the 3,852 corpus sentences of Asai Take's (浅井タケ) west-coast tales (Murasaki 2001).

24.1 Topic marking: *neampe* and *anahne*

The topic particle of Sakhalin Ainu is *neampe* 'as for' (east-coast sources *neanpe*), placed after the phrase it marks. Chiri lists it as the Sakhalin equivalent of Hokkaidō *anak*, *anakne* and of Japanese は (Chiri 1942: §127), and Dal Corso calls it simply the topic marker (Dal Corso 2024: 61). Formally it is the anaphoric pronoun *neampe* 'that thing' — *ne-an-pe* 'thing being that' (Chapter 9) — frozen into a particle, and the pronominal and particle uses still shade into one another. It is by far the commonest information-structure word of the corpus: 917 tokens in the Asai Take tales by our count, against 312 of *ka* and 58 of *kayki*.

With a noun phrase host, *neampe* typically marks a contrastive or switched topic rather than mere givenness. Subjects that simply continue the current topic go unmarked; *neampe* appears where one referent is picked out against another, as when sisters divide their chores or pair off with husbands in sequence — *inoskun monimahpo neampe inoskun horokewpo sam* 'the middle girl married the middle man', and so on down the list (Murasaki (2001) text 18). We analyse this as the particle's core function in the west-coast narrative corpus.

- (1) *kiyanne monimahpo neampe unci-uwaare*
be.eldest girl TOP fire-kindle

'The eldest girl, for her part, kept the fire going.' — Murasaki 2001: text 18; West Sakhalin, Odasu (Asai Take)
The middle and youngest sisters have just been sent out for firewood; *neampe* sets the remaining sister against them.

- (2) *iineahsuy tan horokewpo neampe aynu ka hannehka*
hey this man TOP person even NEG

'Listen — this man is not even a human!' — Murasaki 2001: text 25; West Sakhalin, Odasu (Asai Take)
The speech continues *seta horokewpo nee manu* 'he is a dog-man'. Topic *neampe* and focus *ka* co-occur in one clause.

The same particle serves the east coast. In Yamabe Yasunosuke's (山辺安之助) autobiography *neampe* marks non-subject topics — *urayki neampe ham utara ki kun-pe ne* 'war is a thing one must not make', with the object topicalised and the subject unmarked (Sakaguchi 2020a: 190–191) — and it may be doubled over a contrastive list:

- (3) *ikoro neampe tomi neampe an-i yahka nata koro kun-pe han ne*
treasure TOP riches TOP exist.SG-NMLZ although who have should-NMLZ NEG COP

'Treasures, riches — though they exist, they are not things for anyone to own.' — Piłsudski 1912: text 23, l. 201; Sakaguchi 2020a: 195; East Sakhalin
The speech continues *yaykota e-koro kun-pe ne* 'you must keep it yourself' (Chapter 9, example (7) there).

Beside it the east coast keeps *anahne* 'as for', the formal match of Hokkaidō *anakne*. Piłsudski glosses it 'as regards' in his interlinear English (Piłsudski 1912: 23); in his texts it favours scene-setting adverbs — *tani anahne* 'now, for its part' is a recurrent narrative re-launcher (Chapter 9, example (6) there; Sakaguchi (2020a: 194)). The two markers stack in one sentence:

- (4) *tani anahne anokay neanpe iwan pon sumari uriwahne sumari an-ne-si-te okay-an-ahci*
 now TOP INDEF TOP six small fox sibling fox INDEF.A-COP-PL-and exist.PL-INDEF.S-PL

‘Now we, for our part — we are six little foxes, brother foxes, and we live together.’ — Pilsudski 1912: text 21, ll. 23–25; East Sakhalin, Tunayci (Yasinóske)

A fox in human disguise introduces himself; -si is the east-coast shape of plural -hci.

In the west-coast corpus *anahne* is absent: Asai Take's *anah* (17 tokens, all of them) is the conditional converb ‘if, when’ (Murasaki (2025: 33); Chapter 23; example (19) below), and its combination *anahkayki* with the focus particle *kayki* yields concessive ‘even if, although’. We take the division of labour to be dialectal: the east coast retains the inherited topic particle *anahne* alongside innovating *neanpe*, while the west coast has generalised *neampe* and reserves bare *anah* for the conditional.

Outside the folktales, the expository prose of the Sentoku letters shows the unmarked side of the same system: a topical subject that continues or re-activates an already salient referent is simply placed first, with no particle. The Ainu community is one such standing topic of the third letter, and the clause reporting its response to official pressure fronts it bare:

- (5) *aynu utara tani monasno nu etunne*
 Ainu PL now promptly hear refuse

‘The Ainu now flatly refused to give their assent.’ (lit. ‘the Ainu people now promptly refuse to hear’)

— Tangiku & Ogihara 2001: Letter 3; East Sakhalin, Otasan (Sentoku Tarōji)

The preceding clause is *aynu okta sisan puri ko-nte rusuy yahka* ‘though they want to impose Japanese ways on the Ainu’; *aynu utara* is the continuing topic and stands unmarked, where a switched topic would take *neanpe*.

24.1.1 Clause topics: -hi neampe The majority of *neampe* tokens do not follow a noun at all. Of the 833 tokens in the Asai corpus whose host stands in the same line, 654 by our count follow a clause nominalized with the suffix *-hV* (Chapter 20), yielding a temporal-thematic frame ‘when X happened, ...’ — Dal Corso identifies *-hV* plus *neampe* as a temporal subordinating expression (Dal Corso 2018: 375). We treat both uses as one construction: the particle sets up a frame — a referent or an event — within which the rest of the sentence is interpreted.

- (6) *kinta makan-ih neampe taa sine cise an manu*
 to.mountains go.uphill.SG-NMLZ TOP then one house exist.SG REP

‘When he went up into the mountains, there was a house.’ — Murasaki 2001: text 19; West Sakhalin, Odasu (Asai Take)

Sentence-initially, with no host at all, *neampe* resumes the previous discourse wholesale — ‘as for that, thereupon’ — and functions as a topic-switching connective, frozen above all in *neampe kusu* ‘and so’ and *neampe taa* ‘and then’; this absolute use accounts for most of the remaining tokens. Independent personal pronouns, themselves optional emphatic adjuncts, are the other main carriers of contrast: *kuani* ‘I (and not you)’, *eani* ‘YOU’, frequently reinforced by *neampe* as in *ecokay neanpe cis ohta eci-okay* ‘as for you, you stay in the boat’ (Chapter 9).

24.2 Focus particles

The additive and restrictive particles follow the phrase they scope over, after any postposition. Chiri's inventory for this class gives *ka* ‘also’, *kaiki* ‘also’, Sakhalin *na* ‘also’, *patek* ‘only’ and *poka* ‘at least’ (Chiri 1942: §125); Furukawa's Rayciska grammar lists *ka*, *naa* ‘also’ and *pateh* ‘only’ as the living members of the class (Furukawa 1967: 106).

ka ‘also, even’ is the general additive ‘also, too’, and under negation a scalar ‘(not) even’, as in (2) above and in the negative indefinites *neh ka* ‘nothing at all’, *hemata ka* ‘something/ anything’ built on interrogatives (Chapter 9; Chapter 17). It attaches to nouns, adverbs and converb clauses alike:

- (7) *anokay aynu utara ramma ka pirka ponno ponno cep ukoyki an=e=hci*
 INDEF person PL always also be.good a.little a.little fish catch INDEF.A=eat=PL

‘We Ainu are, as ever, well; we catch and eat just a little fish.’ — Tangiku & Ogihara 2001: Letter 1; East Sakhalin, Otasan (Sentoku Tarōji)

ka on the adverb ramma ‘always’: ‘now as always’.

kayki ‘even, also’ is the emphatic counterpart, ‘even, indeed also’. The two alternate in the same frame within one tale: in text 28 a succession of slighted animals each complain *taah ka ku-e-ocis* ‘this too I resent’, but at the climax the herring-girls say:

- (8) *tah kayki ku-e-ocis kusu ku-an*
 this even 1SG.A-APPL-resent PURP 1SG.S-exist.SG

‘This, of all things, is what I resent.’ — Murasaki 2001: text 28; West Sakhalin, Odasu (Asai Take)

kusu plus the existential is the progressive (Chapter 15): ‘I am resenting precisely this’.

kayki also welds onto subordinators — *anahkayki* ‘even if’, *(ne)yahkayki* ‘although’ (Chapter 23) — and onto *neewa* ‘and’ in *neewa kayki* ‘but, and yet’. The plain additive of the east-coast letters is *na* ‘also’, which Chiri records as a specifically Sakhalin particle ‘also’ (Chiri 1942: §125); Sentoku Tarōji uses it again and again where Asai Take has *naa* (258 tokens, including the correlative *muysankeh ne naa sahka ne naa* ‘both into brooms and into chopsticks’, text 5):

- (9) *cokay na hateki ponno ci=kara*
 1PL.EXCL also field a.little 1PL.EXCL.A=make

‘I too have made a small field.’ — Tangiku & Ogihara 2001: Letter 2; East Sakhalin, Otasan (Sentoku Tarōji)

cokay with singular reference, as regularly in the letters; hateki is Japanese hatake ‘field’.

The restrictive particle is *pateh* ‘only’, underlyingly *patek* with the regular debuccalization of the final stop; the east-coast texts show *pate* ~ *pateki* with the final consonant lost or supported by a vowel (Dal Corso 2024: 61), (Chiri 1942: §125).

- (10) *neeteh tani taa sineh pateh taa an manu*
 and.then now then one only then exist.SG REP

‘And then only one was left.’ — Murasaki 2001: text 37; West Sakhalin, Odasu (Asai Take)

Of the dog-men, all the others have been slain.

- (11) *kesanto nispa pateki tere okay=ahci*
 every.day master only wait exist.PL=PL

‘Every day they wait only for you, sir.’ — Tangiku & Ogihara 2001: Letter 3; East Sakhalin, Otasan (Sentoku Tarōji)

Of the addressee’s Sakhalin family; tere plus the plural existential is progressive.

Two of Chiri’s particles need comment. *poka* ‘at least, even’ carries no Sakhalin tag in his list, and in our Sakhalin corpus every *poka* is the locative postposition ‘over, along’ (*rih poka ahkas cikah* ‘the birds that walk the sky’; Chapter 11): we conclude that the scalar particle is a Hokkaidō item whose work falls in Sakhalin to *kayki* and *ka*. Finally, the pure emphasis particles *asi(n)* and *es* ~ *eski* belong to the literary register of the east-coast tales; Piłsudski judged *es* “an expletive, unmeaning..., used, it seems, only to give animation to the tale” (Piłsudski 1912: 167), and Chiri retained them as Sakhalin emphasis particles (Chiri 1942: §127).

24.3 Word-order variation

The pragmatically neutral clause is verb-final, with topical material first (Chapter 19); interrogatives stand in situ (Chapter 9), so there is no dedicated preverbal focus slot to fill by movement. The living word-order resource of the spoken narrative is instead the postverbal afterthought: a constituent added after the sentence-final verb (usually after the reportative *manu*, Chapter 16) to clarify a referent the speaker has left underspecified. In the Asai Take corpus we count 81 such right-dislocations after *manu* alone, ranging from full noun phrases to place adverbs; the commonest is specification of the subject of a verb of saying.

- (12) *nah taa yee manu acahcipo*
 QUOT then say REP old.woman

‘So she said — the old woman.’ — Murasaki 2001: text 36; West Sakhalin, Odasu (Asai Take)

- (13) *nah ahkapo ahkapo pisi kayki taa ampene taa mokoro- manu tu horokewpo*
 QUOT younger.brother.younger.brother ask even then utterly then sleep-PL REP two man

‘Even when their little brother asked, they slept right on — the two men.’ — Murasaki 2001: text 54; West Sakhalin, Odasu (Asai Take)

The dislocated *tu horokewpo* resolves the plural subject of *mokoro-hci*.

- (14) *amahka-ha-hcin sineh an manu monimahpoo*
 daughter-POSS-PL one exist.SG REP girl

‘They had one daughter — a girl, that is.’ — Murasaki 2001: text 28; West Sakhalin, Odasu (Asai Take)

The preceding sentence is *sine horokewpo macihi tura okayahci manu* ‘a man lived with his wife’; the postposed bare noun glosses the kin term, with the lengthened final vowel of the afterthought contour.

We analyse these as true afterthoughts rather than a focus position: the dislocated phrase is always discourse-old or inferable, never the answer to a question, and the clause is already complete without it. The same looseness lets a frame adverbial trail the verb (*sine cih yan manu, atuy kaawa* ‘a boat came ashore — from the sea’, text 28). Fronting, by contrast, is the normal home of *neampe*-marked topics (§24.1), which precede everything else in their clause.

24.4 Presentational constructions

New referents enter the discourse through an existential presentation: the numeral *sine* ‘one, a certain’ plus noun plus *an* ‘exist.SG’ (or *okay* ‘exist.PL’; Chapter 18). Chiri observed that *sine* here does the work of an indefinite article and the anaphoric demonstratives that of the definite — his model sequence is *sine aynu an. ne aynu ene itaki* ‘a man was. that man spoke thus’ (Chiri 1942: §100, §103). Forty of the fifty-four tales in the Asai Take corpus open with exactly this formula, usually anchored to the stock village name *Sannupis*:

- (15) *sannupis-ta sine horokewpo an manu*
 Sannupis-LOC one man exist.SG REP

‘In Sannupis there lived a man.’ — Murasaki 2001: text 19; West Sakhalin, Odasu (Asai Take)

The tale-opening formula; the next sentences give his habitual life with *koh* ‘whenever’ clauses, and *sine too* ‘one day’ then launches the plot.

A variant places the numeral after the noun in its nominalized form *sineh* (Chapter 10): *ahci sineh an manu* ‘there was an old woman’ (text 1), and example (14) above. Inside a running narrative the presentation is typically framed by a *-hi neampe* clause topic, as in (6): the protagonist’s movement sets the scene, the existential introduces the new participant, and the anaphorics *neya* and *nean* ‘the aforementioned’ then track it for the rest of the tale (Chapter 9).

- (16) *otakaa-ta san-ih neampe taa sine cih yan manu*
 beach-LOC go.down.SG-NMLZ TOP then one boat come.ashore REP

‘When she went down to the beach, a boat came ashore.’ — Murasaki 2001: text 28; West Sakhalin, Odasu (Asai Take)

The girl is the established topic; the boat is presented with *sine* and at once becomes the topic of the next clause.

24.5 Discourse particles in dialogue

Quoted dialogue and the conversational record show a compact system of sentence-final particles. Directives are graded by three request particles, in descending politeness *kanne*, *waa*, *koh* (Murasaki 2025: 38): *kanne* makes a courteous request (*hanka cis kanne* ‘don’t cry, dear’, text 3), *waa* a plain command, and *koh* — Furukawa’s *kah* ‘rough command, used to one’s own children’ (Furukawa 1967: 107) — a curt one (*tah ampa koh!* ‘take this!’). Plural addressees take the imperative suffix *-yan*, combinable with *kanne*:

- (17) *pirikano paye-yan kanne*
 well go.PL-IMP.PL FIN

‘Go carefully, all of you — do.’ — Murasaki 2025: 38; West Sakhalin

- (18) *ayroo ahun-u waa teeta ahun-u waa*
 come.on enter-EP FIN here enter-EP FIN

‘Come now, come in! Come in here!’ — Murasaki 2001: text 3; West Sakhalin, Odasu (Asai Take)

A woman takes in the weeping girl; -u is the epenthetic vowel before the final particle, and *ayroo* the inviting interjection.

The same *wa* closes requests in the letters — *hanka e=ranpotara wa* ‘do not worry’ (Tangiku & Ogihara 2001: Letter 3) — so the particle belonged to both coasts and to written as well as spoken usage. Invitations have polite *anaa* ‘let us...’ and familiar *roo* (*keh, paye-an roo* ‘come on, let’s go!’); assertions directed at the hearer take *naa* ‘you’ll see’ and *wanaa* ‘see, just as I said’; and an indignant refusal ends in *nanko* (Furukawa 1967: 107–108):

- (19) *nah e-kii anah ommo orowa an-ko-caaranke naa*
 thus 2SG.A-do if mother by INDEF.A-APPL-scold FIN

‘If you do that, you’ll get a scolding from your mother, mark my words.’ — Furukawa 1967: 108; West Sakhalin, Rayciska

The indefinite-person verb with an *orowa* agent is the Sakhalin passive (Chapter 14); *naa* stakes the speaker’s confidence on the prediction.

The remaining final particle, *hVV* (*hii, hee...*, its vowel copied from the verb), strengthens statements and questions alike, the two distinguished only by intonation (Murasaki (2025: 19); Chapter 9); Chiri’s bare question particle *he* belongs to the same interjectional class (Chiri 1942: §129). Dialogue turns themselves are managed by interjections, above all *iineahtsuy* ‘hey, say’, which opens an appeal to the hearer many times over in the Asai Take corpus — example (2) — and shades from ‘hey!’ into ‘now listen, the fact is...’. Together with the topic chains of §24.1 and the presentational formula of §24.4, these little words are what give the spoken Sakhalin sentence its characteristic shape: frame first, verb last, and the speaker’s stance hung on the end.

Part VI · The lexicon and language contact

Chapter 25 · Language contact and loanwords

Loanwords from Japanese, the Manchu–Amur trade network, Nivkh, and Russian; their phonological adaptation and morphological integration; and the Cyrillic epistolary register.

Sakhalin was a meeting place. Within reach of a single household the language touched Japanese to the south, the Tungusic of the Uilta (Orok) and the wider Manchu-Amur trade network to the west, the isolate Nivkh (Gilyak) of the central island and the lower Amur, and, from the nineteenth century, Russian. The sociohistorical frame — who spoke what to whom, the interpreters and trade fairs, the partition of 1905 — is set out in Chapter 1 (The language and its speakers), and is not repeated here; nor is the regular adaptation of foreign sounds, whose basics belong to Chapter 4 (Phonology). This chapter takes up what those do not cover: the lexical strata themselves, source language by source language, with their semantic domains and their evidential weight; the most thorough single treatment is Shiraishi and Tangiku's contact chapter in the Bugaeva handbook (Shiraishi & Tangiku 2022), whose loanword appendix underlies much of the tabulation below. The sections run through the Japanese stratum (§25.1), the Manchu-Amur trade stratum that reached Ainu mostly through Tungusic and Nivkh middlemen (§25.2), the large and under-studied Nivkh layer (§25.3), the thin Russian lexicon and the Cyrillic writing register that is its real legacy (§25.4), the phonological adaptation of loans (§25.5), and their morphological integration (§25.6). A word of caution governs the whole: borrowing direction on Sakhalin is often genuinely undecidable, and several items long treated as loans into Ainu are now thought to run the other way. Where the evidence is unsettled it is flagged as such rather than smoothed over.

25.1 The Japanese stratum

Japonic contact with Ainu is the oldest documented layer and falls into two very unequal parts. A small, deep stratum entered before the island period: items such as *sippo* 'salt' 'salt' (Old Japanese *shiho*) and *pasuy* 'chopsticks' 'chopsticks' preserve features lost in later Japanese — the stop pronunciation of **p*, the diphthong **uy* — and so must have been received around the middle of the first millennium CE, when the ancestor of Ainu still lay in the immediate northern neighbourhood of the expanding Yamato state Janhunen (2022). These are pan-Ainu and not specifically Sakhalin; *sippo* is recorded in the earliest Hokkaidō and Kuril word-lists alike.

The bulk of the Japanese vocabulary is far younger and belongs to the Karafuto era of the Russo-Japanese frontier and the Japanese administration after 1905, when material culture flowed north in quantity. The diagnostic Sakhalin trait here is prosodic: the pitch accent of the Japanese source is rendered as *vowel length*, the Sakhalin counterpart of the Hokkaidō pitch contrast (Chapter 4 (Phonology)), so that an accented syllable surfaces long. The clearest set is the 'metal' family, all built on *kaani* 'metal' (< Japanese *kane*, with the long vowel for the accent), beside Hokkaidō *kani* with a short vowel Hattori (1964: 261).

Sakhalin	gloss	Japanese source	note / Hokkaidō
<i>kaani</i>	metal, iron	kane 金	length for accent; [HA] <i>kani</i> (Hattori 1964: 261)
<i>konkaani</i>	gold	kogane / konkane 金	(Hattori 1964: 261)
<i>sirokaani</i>	silver	shirogane 銀	(Hattori 1964: 261)
<i>tuuki</i>	cup (for liquor)	tuki 坏	length; [HA] <i>tuki</i> (Hattori 1964: 117)
<i>tehpō</i>	gun	teppō 鉄砲	geminate → <i>hp</i> ; [HA] <i>teppo</i> (Hattori 1964: 118); (Shiraishi & Tangiku 2022) (no. 74)
<i>peko</i>	cow	beko	initial <i>b</i> → <i>p</i> (§25.5)
<i>puta</i>	pig	buta 豚	initial <i>b</i> → <i>p</i>
<i>tampaku</i>	tobacco	tabako 煙草	also via the trade network; see §25.2
<i>kanpi</i>	paper, letter	kami 紙	pervasive in Sentoku's letters; cf. (1)

Table 42: Karafuto-era Japanese loans in Sakhalin Ainu. Forms and the Hokkaidō comparison are from Hattori's dialect dictionary (1964); the Sakhalin shape differs from Hokkaidō chiefly in vowel length and in coda repair (the row for 'gun'). The gloss '[HA]' marks the Hokkaidō form given for contrast.

Two of these words carry the chapter's recurring lesson about direction and route. *tampaku* 'tobacco' is formally a Japanese loan (*tabako*), but the same item travels the whole northern network — Nivkh *tamχ*, Uilta *saŋna*, Manchu *dambaku* (Shiraishi & Tangiku 2022) — so a Sakhalin speaker could have met it from more than one side. *kanpi* is the workhorse of the only substantial native-written Ainu prose we have, the Sentoku letters of 1906, where it means both the physical 'paper' and the 'letter' written on it; its first attested clause already turns on the wartime shortage of Russian paper:

- (1) тани нуця исан ренкайне нуца канбіка ане сирун кусу канби пон چیцяі ани канби аноманде,
tani nuca isan renkayne nuca kanpi ka an=esiriku su kanpi pon cicay ani kanpi an=omante
 now Russ- not.exist be- Russ- paper even INDEF.A black paper small scrap INS letter INDEF.A=send
 ian cause ian cause

'Now that the Russians are gone, we are short even of Russian paper, so we send our letter on a little scrap of paper.' — Tangiku & Ogihara 2001: Letter 1; East Sakhalin, Otasan (Sentoku Tarōji)

Two loans cooccur: *kanpi* 'paper, letter' (< Japanese *kami*) and *nuca* 'Russian' (not from Russian directly; see §25.2). The clause is glossed in full in Chapter 26 (Glossed texts), §sentoku-letter.

Spoken code-switching into Japanese is also recorded, increasingly so in the latest generation of speakers. In Asai Take's post-war west-coast recordings, Japanese asides and discourse words are dropped into an otherwise Ainu grammatical frame — the narrator breaks into Japanese for an evaluative comment and returns to Ainu for the event line (Murasaki 2001). We treat such material as switching rather than borrowing, since the inserted Japanese keeps its own phonology and is not integrated into Ainu morphology (§25.6).

25.2 The Manchu–Amur trade stratum

For two centuries before the Russian and Japanese states arrived in force, Sakhalin hung off the northern end of the Manchu tributary-trade system that ran up the Amur: marten and sable pelts moved south, and silk, metalware, tea and spirits moved north, passing through Uilta and Nivkh hands before they reached the Ainu. The vocabulary of this trade is the clearest case in the language of borrowing *by relay* rather than face to face. Shiraishi and Tangiku's count is blunt about it: of the loanwords shared across the island's three languages, Nivkh and Uilta share by far the most (61 items), Nivkh and Ainu rather fewer (29), and Ainu and Uilta fewest of all (some 13) — a distribution that matches the historical record, in which the Ainu generally did not understand Uilta and communication ran one-way from the Uilta who knew Ainu (Shiraishi & Tangiku 2022). Direct Tungusic-to-Ainu borrowing is therefore thin; most Tungusic-looking words in Ainu arrived through the network, not from the Uilta next door.

The structurally most important members of this stratum are the borrowed counting-bases *tanku* 'hundred' 'hundred' and *wantanku* 'thousand', which match Manchu-Tungusic *tanggū* and underpin the productive decimal numeral system that Sakhalin Ainu developed for trade arithmetic; Pilsudski was told outright that

tanku came ‘from the Olcha’ with the marten trade Pilsudski (1912: 141). Because they are already treated at length under the numerals, they are only cross-referenced here (Chapter 10 (Numerals and quantification)). The remaining trade words cluster in the expected domains – drink, ritual, status, and the exotic goods themselves.

Ainu	gloss	source / route	citation
<i>cay</i>	tea	Manchu <i>cai</i> (N <i>c^hai</i>)	(Shiraishi & Tangiku 2022) (no. 6)
<i>arakke</i>	spirits, liquor	Manchu <i>arki</i> , ultimately Arabic ‘ <i>araq</i> ’	(Tamura 1999); (Shiraishi & Tangiku 2022) (no. 8)
<i>tonkori</i>	five-string zither	Manchu <i>tenggeri</i> (N <i>tiŋriŋ</i>)	(Shiraishi & Tangiku 2022) (no. 42)
<i>canki</i>	official	‘Santan’ (Amur) trade jargon	(Chiri 1942: §17(2))
<i>nuca</i>	Russian (person)	Orok/Uilta <i>luča</i>	(Chiri 1942: §17(2))
<i>kaco</i>	shaman’s drum	Gilyak (Nivkh) <i>kaš</i>	(Chiri 1942: §17(2))

Table 43: Trade-stratum loans reaching Ainu through the Manchu-Amur network. ‘M’ marks a Manchu comparandum, ‘N’ a Nivkh intermediary, where the source specifies one. Etymologies after Chiri (1942) and Shiraishi & Tangiku (2022); item numbers in the latter’s appendix are given for the cross-language sets.

The word *nuca* ‘Russian’ deserves a note because it is so often misattributed. It is *not* a borrowing from Russian; it comes from Orok/Uilta *luča*, with the initial *l* resolved to *n* by the Ainu avoidance of liquid onsets (§25.5), so that the name the Ainu used for the Russians reached them through their Tungusic neighbours rather than from the Russians themselves Chiri (1942: §17(2)). It is one of the highest-frequency loans in the corpus: every one of Sentoku’s three letters turns on the comings and goings of the *nuca* (see (1)). The zither name *tonkori* is a second cautionary case – the instrument the Nivkh call *tiŋriŋ* is organologically quite unlike the Ainu *tonkori*, so the *name* spread along the trade route even though the *thing* did not, and the two should not be equated (Shiraishi & Tangiku 2022). Finally, the west-coast mitten word *wampakka* ‘mittens’ is a genuine direct Uilta loan (Uilta *wambakka*), but a strictly local one: it is a west-coast isogloss against east-coast native *matumere*, treated under the dialects (Chapter 2 (Dialects, sources, and previous research)).

25.3 The Nivkh layer

The contact layer most likely to be under-represented in earlier grammars is the Nivkh one. Where the present description’s predecessors cite barely a word or two, Shiraishi and Tangiku’s appendix lists around twenty-nine items shared between Nivkh and Ainu, concentrated in the technologies of the central island – dog-sledding, sea-hunting, containers, and the fauna and tackle of the coast (Shiraishi & Tangiku 2022). The geography is telling: the Nivkh dialect richest in shared vocabulary is that of Poronaïsk, in the central contact zone where Ainu, Nivkh and Uilta territories met, which marks the borrowing as local rather than island-wide.

Ainu	gloss	Nivkh	appendix no.
<i>nuso ~ nuhsu</i>	dog sled (the whole rig)	<i>nuci</i>	1
<i>mahru</i>	towline spreader	<i>maxt</i>	16
<i>mohrasi</i>	sled runner	<i>motas</i>	9
<i>muhru</i>	pillow	<i>mut</i>	35
<i>herohki</i>	Pacific herring	<i>heruk</i>	21
<i>sintoko</i>	barrel, cask	<i>sintux</i>	18
<i>kaya</i>	fish-skin garment / sail	<i>qai</i>	48
<i>yuuru</i>	bow of a spring trap	<i>jur</i>	38
<i>konko</i>	bell	<i>qong</i>	50
<i>nikari</i>	ladder, stairs	<i>ɲɲar</i>	49
<i>kaco</i>	shaman's drum	<i>q^has</i>	39
<i>tonto</i> †	sea-lion / seal hide	<i>tun</i>	25
<i>upun</i> †	outdoor winter storehouse	<i>uvɲ</i>	47

Table 44: Nivkh–Ainu shared vocabulary (Shiraishi & Tangiku 2022, with their appendix numbers). Many of these show the regular Nivkh *t* ↔ Ainu *r* coda correspondence discussed in §25.5. The final group is flagged for direction: these are likelier to be Ainu → Nivkh.

The dagger marks the candour the data demand. For the dog-sled word, Chiri held *nuso* to be a borrowing from Nivkh *nuci* – there is no native Ainu word for the rig, and the *nuci* set is spread across the lower Amur – and Shiraishi and Tangiku accept this (Shiraishi & Tangiku 2022). But for *tonto* ‘tanned hide’ and *upun* ‘storehouse’ the likelier direction is the reverse, Ainu → Nivkh, so they are best not counted as Ainu loans at all. The general point is that for a great many of these the direction is simply debated; some may even come from a third, lost source – the language of the Okhotsk culture – that fed both Ainu and Nivkh Janhunen (2022). The reindeer word is the most famous instance of an old, deep exchange in this zone: Ainu *tunakay* ‘reindeer’ ‘reindeer’ goes back to Pre-Proto-Gilyak **tola-ŋa.y* ‘pulling animal’, and was in turn passed on into Japanese as *tonakai* – Ainu here standing as the relay between Nivkh and Japanese, the mirror image of its receiving role elsewhere in the chapter Janhunen (2022), (Austerlitz 1976).

25.4 Russian and the Cyrillic register

Russian left Sakhalin Ainu almost no integrated lexicon. Such direct loans as are claimed are few and uncertain – *kosika* ‘cat’ (< *koshka*) is the stock example, but it is poorly attested and we flag it as unverified. The Russian contribution to the record is of a different order: it is a *register* and a *script*, not a word-list. Two things stand out. The first is the Cyrillic writing tradition. Bronisław Piłsudski ran a short-lived east-coast school at which Ainu was written in the Russian alphabet, and the most substantial surviving body of Ainu prose by a native hand is the set of three letters that Sentoku Tarōji wrote him in Cyrillic in 1906, edited by Tangiku and Ogihara (Tangiku & Ogihara 2001). The orthographic side of this – how Cyrillic was made to register Ainu, and what it could not register (vowel length above all) – belongs to the orthography chapter (Chapter 3 (Orthography and transcription)), and the letters are glossed continuously in the texts chapter (Chapter 26 (Glossed texts)). The second is the epistolary register the letters establish: an Ainu opening formula *ituyaskara nispa* ‘beloved master’ over a frame studded with Russian apparatus – calendar dates and place phrases (*4go Fevral’* ‘4th of February’, dating clauses on the Russian model), ruble-and-kopeck sums, and Russian personal and place names adapted to Ainu phonology: Korsakov, Aleksandr → *Ariksandoru*, *lesnichiy* ‘forester’ → *Lesneci*. These are best read not as loans into the language but as the written furniture of a bilingual letter-writer; several further items that appear in the letters – *atarimay*, *arenda* ‘lease’, *starosta* ‘village elder’, *menco*, *comen* – we treat as unintegrated code-switches into Russian (or, for *atarimay*, Japanese) rather than established borrowings, and mark them accordingly. The point of substance is that Russian influence on Sakhalin Ainu is concentrated in this written, formulaic, name-and-number register, and barely touches the core lexicon at all.

25.5 Phonological adaptation of loans

Loanwords are reshaped to Ainu phonotactics by a small set of regular operations, most of them already visible in the tables above. They are synthesised here; the native phonological constraints they answer to are set out in Chapter 4 (Phonology), and only the contact-specific patterns are drawn together.

Liquid-onset avoidance. Ainu has no syllable-initial *l*, and a foreign initial liquid is repaired. The diagnostic case is *nuca* ‘Russian’ < Uilta *luča*, where *l* becomes *n* Chiri (1942: §17(2)). **Accent → length.** The pitch accent of a Japanese source is rendered as Sakhalin vowel length, the systematic correspondence behind *kaani*, *tuuki* against Hokkaidō short-vowel *kani*, *tuki* Hattori (1964: 117, 261). **Voicing neutralisation.** Ainu lacks a voicing contrast in stops, so voiced onsets of Japanese loans surface voiceless: *beko* → *peko* ‘cow’, *buta* → *puta* ‘pig’.

Geminate and cluster repair. A foreign geminate or illicit cluster is resolved into a permitted coda-plus-onset sequence, the coda surfacing as *h* by the regular coda neutralisation of the native phonology: thus Japanese *teppō* → *tehpō* ‘gun’, with *hp* for *pp* Hattori (1964: 118). Two of Chiri’s showcase adaptations carry the repair further, through whole-segment loss and metathesis. The headwear word *atampusu* ‘a kind of hood’ comes from Japanese *watabōshi* 綿帽子 by way of an intermediate *matampusi*, then loss of the initial *m* Chiri (1942: §17(2); §19); and the lamp word, Hokkaidō *ratcaku* < Chinese 蠟燭 (via Japanese), appears in Sakhalin as *cahraku* ‘light, lamp’ by metathesis of the initial *r* Chiri (1942: §20).

The Nivkh *t* ↔ *r* correspondence. Across the Nivkh layer a Nivkh syllable-final *t* answers to an Ainu *r*: *mut* → *muhru* ‘pillow’, *maxt* → *mahru* ‘towline spreader’, *motas* → *mohrasi* ‘sled runner’ (Shiraishi & Tangiku 2022). The correspondence is regular enough to serve as a diagnostic of the layer, though it does not by itself fix the direction of any individual borrowing.

25.6 Morphological integration

The evidence for how loanwords behave in the grammar is thin, and we state only what the texts attest. The clearest fact is that integrated loan nouns are ordinary nouns: they take the native counting frames of Chapter 10 (Numerals and quantification) and slot into the same syntactic positions as inherited vocabulary. In the Sentoku letters the loan *kanpi* ‘letter’ is counted, possessed, and governed by postpositions exactly as a native noun would be, and it heads relative clauses (*e=koro kanpi* ‘your letter’ with the regular second-person possessive prefix). Trade-stratum nouns likewise take native quantifier and classifier frames: the borrowed counting-bases *tanku* and *wantanku* are not merely loaned words but loaned grammatical *nouns*, themselves countable with the native adnominal numerals (*sine tanku* ‘one hundred’), which is as deep an integration as a numeral system admits (Chapter 10 (Numerals and quantification)).

Whether loan *verbs* were taken in is harder to show. The integrated lexicon is overwhelmingly nominal — names of goods, animals, tools and materials — as is typical of trade borrowing, and the corpus offers no secure example of a borrowed verb stem inflecting with the Ainu personal prefixes. Where a foreign action concept is needed, the language reaches instead for a native verb or a light-verb construction with *kara* ‘make’ or *ki* ‘do’, leaving the loan in nominal shape. We flag this as an argument from silence rather than a demonstrated gap: the recorded corpus is small, and the latest speakers’ free code-switching into Japanese (§25.1) shows that a foreign verb could always be reached for by simply switching language rather than by morphologically nativising the stem — which may be part of why a nativised loan-verb class never had to develop.

Part VII · Texts

Chapter 26 · Glossed texts

Fully glossed Sakhalin Ainu texts: a *tuytah* told by Asai Take, a letter by Sentoku Tarōji, and an extract from Pilsudski's materials.

This chapter presents four short Sakhalin Ainu texts in full interlinear glossing. They were chosen to cover the main kinds of record on which this grammar rests: an oral folktale (*tuytah* 'folktale') told on the west coast in 1990 by the last native speaker, Asai Take; a private letter written on the east coast in 1906 by a literate native speaker, Sentoku Tarōji, in Cyrillic script; a dictated folktale (*ucaskuma* 'tale, tradition') taken down on the east coast in 1903 by Bronisław Pilsudski; and two conversational texts from the west-central coast, printed in 1875 in the appendix to Dobrotvorsky's dictionary — the oldest connected Sakhalin Ainu prose. Between them the texts span a hundred and fifteen years, both coasts, the spoken and the written language, and four genres with distinct grammatical profiles. Each text is given with a philological introduction stating what is known of its recording — medium, date, place, narrator, and the chain of transmission from field record to printed edition — followed by consecutive numbered examples (one sentence per example, nothing reordered) and a commentary that points from the text to the chapters where each phenomenon is treated.

The conventions are those of the rest of the grammar. The morphemic line is retranscribed into our romanisation (Chapter 3): the glottal-stop apostrophes of Murasaki's transcription are dropped, and the orthography of each older source — Sentoku's Cyrillic, Pilsudski's diacritics, Dobrotvorsky's Cyrillic as romanised by Sakaguchi — is interpreted as set out in the headnote to its text (§§26.2–26.4 below); where the original spelling is itself of interest it is shown as a separate line above the gloss. Glossing follows the conventions stated in Chapter 13 and the front matter: the indefinite set *an-/an/i-* is INDEF, the reportative *manu* is REP, the quotative *nah* is QUOT, and the pervasive narrative particle *taa* is glossed EMPH. Translations are our own, made against the Japanese of Murasaki (2001), Tangiku & Ogiwara (2001) and Sakaguchi (2021) and the English of Pilsudski (1912).

26.1 A *tuytah* of Asai Take: the *imu* woman

Asai Take (浅井タケ, died 1994), of Odasu on the west coast, was the last native speaker of Sakhalin Ainu (Dal Corso 2024: 58). Between 1984 and the early 1990s Murasaki Kyōko (村崎恭子) recorded from her, in tape-recorded sessions, several dozen *tuytah*, which Murasaki transcribed and published in full with Japanese translation, in revised edition, as Murasaki (2001). The publication chain is thus short and modern — speaker to tape to editor — and the transcript preserves the texture of live telling: hesitations, clipped variants, and interviewer turns. Murasaki's transcription is broad; its apostrophes mark the predictable glottal onset of vowel-initial words, and we drop them in retranscription, since the onset is not contrastive (Chapter 4, Chapter 3). The extract below is the opening of text 40, *imuu monimahpo* 'The *imu* girl', recorded on 7 April 1990. The thirteen sentences are consecutive and complete — this opening happens to be free of the Japanese asides and interviewer turns that intersperse many of the tales — and they carry the story to its first turning point: immediately after (13) the *imu* woman breaks into the nonsense verse (*niiteh kaa cira rii too...*) that is the set piece of the tale. *imu* is the dissociative startle behaviour, well known across Ainu communities, in which the affected person (usually a woman) echoes words or shouts formulaic verse when startled; here it is played for comedy.

- (1) *orowa sine kotan-ta sine horokewpo yay-reske manu*
 then one village-LOC one young.man REFL-raise REP

'Then: in a village, they say, a young man lived all alone.' — Murasaki 2001: text 40; West Sakhalin, Odasu (Asai Take)

yay-reske 'raise oneself' is the conventional opening formula for a protagonist who grew up, and lives, by himself.

- (2) *sine horokewpika taa ni- koh ku- ceh- koh ee nah kii yayntaa sine too taa kintani- kusumakamanu*
 one young.male and EMPH gather.firewood burn fish- ITR eat QUOT Ho while EMPH one day EMPH mount.gather.firewood good bill SG
 raise do

‘Living by himself, he would gather firewood and burn it, catch fish and eat them; and as he was living so, one day he went up into the mountains to gather firewood.’ — Murasaki 2001: text 40; West Sakhalin, Odasu (Asai Take)

kinta is kim ‘mountain’ + locative -ta with nasal assimilation; the koh ... koh ... nah kii chain is the stock habitual formula of the genre.

- (3) *niina kusu makan- neampe taa sine acahcipo taa niina kusu an manu*
 gather.firewood DURP ihi go.uphill.SG-when EMPH one old.woman EMPH gather.firewood DURP exist.SG REP
 NMLZ

‘When he went up to gather firewood, an old woman was there gathering firewood, they say.’ — Murasaki 2001: text 40; West Sakhalin, Odasu (Asai Take)

We segment makan-*ihi* as makan plus the possessive suffix -hi with epenthetic i, in its clause-nominalizing use: ‘(at) his going up’. The following neampe (ne-an-pe ‘thing which is’) supplies the temporal reading, a stock protasis of west-coast narrative; the second niina kusu an is the purposive read progressively, ‘was (there) in order to gather’ — i.e. ‘was gathering’.

- (4) *neya acahcipo yee manu*
 that old.woman say REP

‘That old woman said, they say.’ — Murasaki 2001: text 40; West Sakhalin, Odasu (Asai Take)

neya ‘that’ tracks the woman just introduced in (3). yee manu opens the quotation, and (8) nah taa yee manu closes it: quoted speech in this corpus is regularly bracketed on both sides, the closing half carrying the quotative nah.

- (5) *iineah- horokewpo an-cise- sanu waa kuani sineh imuu mon- tura ku-an*
 hey young.man ta INDEF.A- go.down.SG FIN 1SG one have.imu girl with 1SG.S- exist.SG
 house- LOC

‘Say, young man, come down to my house! I live with a girl who has imu fits.’ — Murasaki 2001: text 40; West Sakhalin, Odasu (Asai Take)

an-cise-ta ‘at my house’ has the indefinite A-prefix as possessor index; sanu shows a paragogic vowel before the final particle waa. Inside quoted speech the old woman uses the conversational KU-series.

- (6) *imuu mon- tura ku- ku- koh ku-hosini ohta taa niske-ko-ocivwa ohta taa reekohe-imuu*
 have.imu girl with ike ina 1SG.S- 1SG.S- ITR 1SG.S- exist.SG at EMPH load- APPL- exist.SG at EMPH very APPL-
 exist.SG gather.firewood return.SG POSS throw.down have.imu
 and

‘Living with the imu girl, whenever I gather firewood and come home, the moment I throw my load down she has a violent imu fit.’ — Murasaki 2001: text 40; West Sakhalin, Odasu (Asai Take)

V-an ohta, literally ‘at the being of having V-ed’, is a temporal nominalization ‘when (I) V’; niske-he ‘her/my load’ carries the possessive suffix.

- (7) *e-kiororo-an manu kiroro-an manu*
 APPL-excitement-exist.SG REP excitement-exist.SG REP

‘It is most amusing, quite a spectacle.’ — Murasaki 2001: text 40; West Sakhalin, Odasu (Asai Take)

kiororo ~ kiororo ‘force, excitement’; kiroro-an, literally ‘there is excitement’, is ‘to be thrilling, fun’. The reportative manu intrudes into the quotation — the narrator’s voice surfacing inside the character’s, a regular trait of Asai Take’s quoted dialogue.

- (8) *nah taa yee manu*
 QUOT EMPH say REP

‘So she said, they say.’ — Murasaki 2001: text 40; West Sakhalin, Odasu (Asai Take)

- (9) *neeteh tani taa imi-yehe naa neya horokewpo mii-re horokewpo imi-yehe anihi mii*
 and.then now EMPH garment- also that young.man wear- young.man garment- 3SG wear
 POSS CAUS POSS

‘And then she dressed the young man in her own garment, and the young man’s garment she put on herself.’ — Murasaki 2001: text 40; West Sakhalin, Odasu (Asai Take)

- (10) *neetehani taa horokewponii acah-niinaike taa tah taa see teh orowdaa acah-cise- san manu*
 and.then now EMPH young.man EMPH wood cipo gather firewood EMPH this EMPH carry on back then EMPH old woman cipo ta
 EMPH wood old.woman EMPH firewood EMPH this EMPH carry on back then EMPH old woman house go down REG
 LOC

‘And then the young man took on his back the wood that the old woman had gathered, and went down to the old woman’s house, they say.’ — Murasaki 2001: text 40; West Sakhalin, Odasu (Asai Take)

nii acahcipo niina ‘the wood the old woman had gathered’ is a gapped relative clause with no relativizer. The source prints achipo here for acahcipo, one of Asai Take’s frequent short forms.

- (11) *neya monimahpo sinenehka an kusu an manu*
 that girl alone exist.SG PURP exist.SG REP

‘That girl was at home all alone, they say.’ — Murasaki 2001: text 40; West Sakhalin, Odasu (Asai Take)

an kusu an, literally ‘is so as to be’, is the durative periphrasis built on the purposive — the same construction as niina kusu an in (3); sinenehka ‘all alone’ picks up the yay-reske ‘raised himself’ of the opening, the tale’s two solitaires now in parallel.

- (12) *neyke taa nii taa cise apa caa ene ociwe*
 and.then EMPH wood EMPH house door edge to throw.down

‘And then he threw the wood down at the doorway of the house.’ — Murasaki 2001: text 40; West Sakhalin, Odasu (Asai Take)

cise apa caa ‘the edge of the house door’ is a chain of bare juxtaposed nouns headed by the relational noun caa ‘edge, mouth’, with the allative ene; the bare ociwe here against ko-ociwe in (6) shows the applicative doing real work — the old woman’s ko-ociwe ‘throw down at (the house)’ indexes the goal that (12) spells out as a full postpositional phrase.

- (13) *neyke taa cise onnayke-ta imuu hawe-he an manu*
 and.then EMPH house inside-LOC imu voice-POSS exist.SG REP

‘And then, from inside the house, came the voice of her imu fit, they say.’ — Murasaki 2001: text 40; West Sakhalin, Odasu (Asai Take)

onnayke ‘inside’ is a relational noun taking the locative -ta; imuu hawe-he an ‘the voice of (her) imu was’ has hawe ‘voice’ in its literal, fully nominal use, with the possessive suffix — the same noun that, bleached, serves the hearsay evidential construction.

The extract is a compact display of the grammar of west-coast narrative. Nearly every narrative sentence closes with the reportative *manu*, the genre signature of the *tuytah* (Chapter 16), while quoted dialogue inside the tale switches to the unmarked conversational register: the old woman refers to herself with the KU-series (*ku-an*, *ku-niina*, *ku-hosipi*), and narrative third persons are zero-marked throughout (Chapter 13). The habitual chain *niina koh kuruu*, *ceh-kii koh ee* with iterative *koh*, and the progressive *an kusu an* of (11), belong to Chapter 15; the clause-linkers *teh*, *ike*, *neeteh*, *neyke* and the temporal nominalizations *makan-ih* *neampe* and *V-an ohta* to Chapter 23 and Chapter 20. Possessive morphology is dense even in this short stretch (*imi-yehe*, *niske-he*, *hawe-he*; Chapter 7), and the valency machinery of Chapter 14 appears in *yay-reske*, causative *mii-re* ‘dress someone’, and applicative *e-imuu* ‘have an imu fit at (something)’. The assimilation in *kinta* is treated in Chapter 5.

26.2 A letter of Sentoku Tarōji (1906)

Sentoku Tarōji (千徳太郎治, 1872–1929), of Otasan on the east coast, was Pilsudski’s interpreter and collaborator and later the author of a Japanese-language account of his people (Sentoku 1929). In 1906, after Pilsudski had left Sakhalin, Sentoku wrote him three letters in Ainu from Naybuci, using the Russian Cyrillic alphabet — the most substantial body of Ainu prose written by a native speaker in Cyrillic script (Chapter 3). They were deciphered and published, with Latin transliteration and Japanese translation, by Tangiku Itsuji (丹菊逸治) and Ogihara Shinko (荻原眞子) (Tangiku & Ogihara 2001). We gloss the opening of the first letter, dated 4 June 1906. The manuscript is written in narrow physical lines that cut across words and sentences; we have

rejoined them, resolving the hyphenations, and we give the Cyrillic original for the first six sentences. One opaque clause in the long final period is omitted at a clause boundary, marked by an ellipsis; nothing else is left out, and no word is altered. The transliteration follows Tangiku and Ogihara, including their clitic notation *an=*, *ci=*, *e=*.

The script sets the limits of the evidence. Russian Cyrillic registers the consonant skeleton of Ainu well, but it has no device for vowel length, and unstressed vowels waver in Sentoku's spelling: цкін for *cikin*, кестоно for *kestoono*, піка for *pirka*. Quantity and some vowel qualities are therefore restored editorially, while everything carried by consonants — above all the person clitics — can be read directly off the page. Forms whose vowels matter are flagged in the notes.

(14) *итуйаськара ниспа*

ituyaskara nispa
beloved master

‘Beloved master,’ — Tangiku & Ogihara 2001: Letter 1; East Sakhalin, Otasan (Sentoku Tarōji)

We take *ituyaskara* to contain *tuyaskarap* ‘to hold dear, pity’ with the indefinite object prefix *i-*, ‘one whom we hold dear’; the closing formula of the letter has reciprocal *u-* instead: *nispa utuyas(kara)* ‘master with whom we hold each other dear’.

(15) *сонко аноманде цкін нукараконде ва,*

sonko an=omante cikin nukara konte wa
message INDEF.A=send if look give FIN

‘We send word; when it reaches you, please read it.’ — Tangiku & Ogihara 2001: Letter 1; East Sakhalin, Otasan (Sentoku Tarōji)

omante is *oman-te* ‘cause to go’; *konte* ‘give’ adds a benefactive-imperative nuance to *nukara*, softened by final *wa*.

(16) *тани нуча исан ренкайне нуца канбіка ане сирун кусу канби пон чіцяі ани канби аноманде,*

tani nuca isan renkayne nuca kanpi ka an=esirun kusu kanpi pon cicay ani kanpi an=omante
now Russ- not.exist be- Russ- paper even INDEF.A=black paper small scrap INS letter INDEF.A=send
ian cause ian cause

‘Now that the Russians are gone, we are short even of Russian paper, so we send our letter on a little scrap of paper.’ — Tangiku & Ogihara 2001: Letter 1; East Sakhalin, Otasan (Sentoku Tarōji)

We analyse *an=esirun* as applicative *e-* on *sirun* ‘be poor, wretched’ (cf. *sirun kur* ‘poor man’), ‘to be badly off for’; *cicay* ‘scrap’ is obscure — Tangiku and Ogihara render *pon cicay* as ‘a small scrap’.

(17) *анкоро ниспа темана ан okay экиххетанея?*

an=koro nispa temana an okay e=ki-hi hetane ya
INDEF.A=have master how exist.SG exist.PL 2SG.A=do-POSS Q Q

‘Our master, how are you faring?’ (lit. ‘the living that you lead — of what sort is it?’) — Tangiku & Ogihara 2001: Letter 1; East Sakhalin, Otasan (Sentoku Tarōji)

temana an okay e=kihi is literally ‘the how-being living that you do’, with *okay* nominalized; *hetane* (*he-ta-ne*) is a fused interrogative tag stacked with *ya*.

(18) *анокай айну утара рамака піка поно поно цеп укойки анехці.*

anokay ainu utara ramma ka pirka ponno ponno cep ukoyki an=e-hci
INDEF Ainu PL always too be.well little little fish catch INDEF.A=eat-PL

‘We Ainu are well as ever; we catch a little fish and eat it.’ — Tangiku & Ogihara 2001: Letter 1; East Sakhalin, Otasan (Sentoku Tarōji)

The spelling is at its loosest here: *ramma ka* ‘always too’ is written solid as *рамака* with the geminate simplified, and *ponno ponno* ‘a little’ as *поно поно*; the transliteration restores both, after Tangiku and Ogihara. *an=e-hci* ‘we eat’ carries the plural *-hci* agreeing with the plural eaters.

(19) *кестоно ниспа оровано камбі цока цнукара сонно ияі райкере*

kestoono nispa orowano kampi coka ci=nukara sonno iyayraykere
the.other.day master from letter 1PL.EXCL 1PL.EXCL.A=see truly thank.you

‘The other day we read the letter from you, master; thank you truly.’ — Tangiku & Ogihara 2001: Letter 1; East Sakhalin, Otasan (Sentoku Tarōji)

kestoono is elsewhere ‘every day’ (*kesto-no*); Tangiku and Ogihara take it here as ‘the other day’, which the context requires. *iyayraykere* ‘thank you’ is historically *i-yay-rayke-re* ‘you nearly make me kill myself’.

- (20) *nani hospisonkacokayomante suyyahkani kampikta itakumonasabokotarakta e=omkusu nah e=ye renkayne*
 at.once return mes- 1PL.EXCL.DES although 2SG let- in word soon 2SG.AVillage to 2SG.S+POSS QUOT 2SG.Abs use
 sage ter

‘We wanted to send a reply at once, but since in the words of your letter you said that you would soon be going back to your own village, ...’ — Tangiku & Ogihara 2001: Letter 1; East Sakhalin, Otasan (Sentoku Tarōji)
 hospi = hosipi ‘return’; iani is the letters’ spelling of the 2SG pronoun eani. The omitted continuation — ‘although we sent a letter, the master would already have gone, so we did not send it’ — contains the unclear word pokari and the analytic negative hanne.

- (21) ... *yoo-e=kosisamktacokayuu ciki naa iani e=ammaneynenkaokuyironan tu- kanphonasmanispoktaisora ci=kara*
 ... later 2SG.A+POSS 1PL.EXCL if still 2SG 2SG.S+POSS 1PL.EXCL 1PL.EXCL exist. 2SG let- quickly item as- at send- 1PL.EXCL.A=make
 ese cause ter ter ing

‘... but when we asked later at the Japanese near you and heard that you were still there, we joyfully wrote a letter at once and dispatched it to you, master.’ — Tangiku & Ogihara 2001: Letter 1; East Sakhalin, Otasan (Sentoku Tarōji)
 kirora an turano is literally ‘together with there being joy’; isora ci=kara ‘we made a sending’ is Tangiku and Ogihara’s conjectured reading, supported by e=isira ‘you sent’ in the second letter. The reportative manuy marks the news as hearsay.

The letter shows the written vernacular of an east-coast speaker with no model to imitate, and its grammar is continuous with the spoken corpus. First-person reference alternates between the AN-series (*an=omante*, *an=e-hci*, the address form *an=koro nispa* ‘our master’) and the CI-series (*cokay ~ coka*, *ci=nukara*, *ci=nu*), each clause keeping to one series as the personal agreement rule requires; the KU-series never appears in the letters (Chapter 13). Clauses are chained with *renkayne* ‘because’ and conditional *ciki(n)* (Chapter 23); existential negation is *isan* (Chapter 18, Chapter 17); and reported information is marked with *manuy*, the letters’ form of the reportative (Chapter 16). The Cyrillic spelling is itself linguistic evidence: *ниспа* with palatalized *сь* for the [ɕ]-coloured sibilant, and *канби* beside *kanpi* showing the intervocalic voicing of plosives that Piłsudski also recorded on the east coast (Chapter 4; (Dal Corso 2024: 61)).

26.3 From Piłsudski’s Materials: the woman of Paratunnay

Piłsudski’s *Materials for the Study of the Ainu Language and Folklore* (Piłsudski 1912) contains twenty-seven *ucaskuma* dictated by east-coast speakers in 1903–04, the only substantial record of the eastern dialects (Dal Corso 2024: 54–58). The medium is dictation: the texts were taken down in handwriting, sentence by sentence, at a pace the speaker had to hold to — a circumstance with grammatical consequences taken up at the end of this section (Dal Corso 2024: 59). The collection was published in Cracow in 1912 and is most accessible today in the annotated reprint in Piłsudski’s collected works (Piłsudski 1998). The first two texts carry Piłsudski’s own word-for-word English translation beside the Ainu. Our extract is the opening of text 2 (lines 1–34, p. 45), dictated in January 1903 by Sisratoka, aged 28, of Tarayka on the Bay of Patience. Sisratoka is no anonymous informant: three years later Sentoku’s second letter reports the death of his wife and his move within Tarayka (Tangiku & Ogihara 2001: Letter 2) — the two texts of this chapter document one small community. The passage itself, a sober account of the depopulation of the village of Paratunnay, is as much historical as literary testimony (Chapter 1).

Each example shows Piłsudski’s transcription (accents and all) above our retranscription. The interpretation of his orthography follows Dal Corso (2024: 59–66): *x* is the debuccalized coda plosive, our *h*; *ś* and *ć* are the palatal allophone of /s/ and the affricate /c/; *j* is /y/; and the voiced letters *b d g* note the intervocalic voicing of /p t k/, which is not phonemic and is undone in retranscription (Chapter 4). His acute accents we reproduce in the transcription line but do not interpret: they are left out of the retranscription.

- (22)
- Parátunnaj oxta etókota ájnu poróno án.*

Paratunnay ohta etokota aynu porono an
 Paratunnay in formerly person many exist.SG

‘In Paratunnay there were formerly many people.’ — Pilsudski 1912: text 2, ll. 1–3; East Sakhalin, Tarayka (Sisratoka)
 The opening line shows the orthographic system at a glance: oxta with x for the debuccalized coda (our ohta), j for /y/ in Parátunnaj, and the accents we leave uninterpreted. porono an ‘were many’ has the quantifier verb-adjacent, the existential construction of quantified NPs.

- (23)
- Emújke rajaxcí.*

emuyke ray-ahci
 all die-PL

‘They all died.’ — Pilsudski 1912: text 2, l. 4; East Sakhalin, Tarayka (Sisratoka)
 The plural suffix -(a)hci, spelt -axcí, shows the [x] realization of the debuccalized plosive before c.

- (24)
- Okáketa šine máxneku, ré Inanupíríka, širánkuri utara tura pate oměka.*

okaketa sine mahneku re Inanupirika sirankuri utara tura pate omeka
 afterwards one woman name Inanupirika kinsman PL with only remain

‘Afterwards one woman, Inanupirika by name, was left with only her kinsfolk.’ — Pilsudski 1912: text 2, ll. 5–10; East Sakhalin, Tarayka (Sisratoka)
 omeka ‘to remain, be left over’ recurs in Sentoku’s first letter (nuca ponno omeka ‘a few Russians remain’); re ‘name’ stands in apposition.

- (25)
- Šine cise pate ájnu páxteno án.*

sine cise pate aynu pahteno an
 one house only person sufficiently exist.SG

‘There were people enough only for a single house.’ — Pilsudski 1912: text 2, ll. 11–13; East Sakhalin, Tarayka (Sisratoka)
 pahteno ‘just enough, in proportion’ is recorded with the same sense in the modern Sakhalin teaching wordlists.

- (26)
- Néte Tarájkaun nišpa Inanupíríka sám rusúi.*

neete Tarayka un nispa Inanupirika sam rusuy
 and.then Tarayka ATTR master Inanupirika marry DES

‘Then a rich man of Tarayka wished to marry Inanupirika.’ — Pilsudski 1912: text 2, ll. 14–16; East Sakhalin, Tarayka (Sisratoka)
 un is the attributive ‘belonging to (a place)’, as in tan cise un monimahpo ‘the girl of this house’.

- (27)
- Sám rusúike, Inanupíríka etúnne.*

sam rusuy-ike Inanupirika etunne
 marry DES-and Inanupirika refuse

‘Though he wished to marry her, Inanupirika would not have him.’ — Pilsudski 1912: text 2, ll. 17–18; East Sakhalin, Tarayka (Sisratoka)
 Pilsudski writes rusúike as one word: the linker -ike fuses with rusuy, whose final glide is absorbed — we segment rusuy-ike on the strength of the bare rusúi of (26). The tail-head repetition sam rusuy ... sam rusuy-ike is a discourse pattern of dictated narrative, the speaker re-anchoring each sentence in the last.

- (28)
- Tarájkaun nišpa kotánu oxta xošíbi hemaka.*

Tarayka un nispa kotan-u ohta hosipi hemaka
 Tarayka ATTR master village-POSS to return.SG finish

‘The rich man of Tarayka went back to his village.’ — Pilsudski 1912: text 2, ll. 19–21; Dal Corso 2024: 66; East Sakhalin, Tarayka (Sisratoka)
 kotan-u ‘his village’ shows the possessive suffix with elided h (kotan-[h]u), a regular east-coast alternation.

(29) *Inanupíríka tani utárhi ájnu éngo súj ráj.*

Inanupíríka tani utar-hi aynu enko suy ray
 Inanupirika now people-POSS person half again die

‘Now, of Inanupirika’s people, half again died.’ — Pilsudski 1912: text 2, ll. 22–25; East Sakhalin, Tarayka (Sisratoka)
 súj for suy ‘again’ shows both palatal *ś* and *j* for /y/ in one syllable, undone in retranscription. *utar-hi aynu enko* ‘half of her people’ is a partitive of bare juxtaposition; bare *ray* here against *ray-ahci* in (23), with plural subjects both times, shows the *-(a)hci* plural as optional rather than obligatory agreement.

(30) *Nax án ani, táta ohórono Parátunnaj-ta án kojákuś.*

nah an ani tata ohorono Paratunnay-ta an koyakus
 QUOT exist.SG because there long.time Paratunnay-LOC exist.SG be.unable

‘This being so, it was impossible to stay long there in Paratunnay.’ — Pilsudski 1912: text 2, ll. 25–29; East Sakhalin, Tarayka (Sisratoka)

nah an ani ‘by its being so’ presses the instrumental *ani* (cf. *cicay ani* ‘with a scrap’ in §26.2) into causal service; the next sentence varies the same formula as *nah an renkayne* — the two causal postpositions interchangeable in the fixed connective.

(31) *Nax án rénkajne, hejáo japaxcí, utárhi tura isínne japaxcí, Moriruesán japaxcí.*

nah an renkayne heyao yapa- utar-hi tura isinne yapa- Morirue- yapa-
hci hci hci san hci
 QUOT exist.SG because seaward go.by.boat- people- with all.together go.by.boat- Morirue- go.by.boat-
 PL POSS PL san PL

‘And so they took to the sea: with all her people together they went off by boat, and made land at Moriruesan.’ — Pilsudski 1912: text 2, ll. 30–34; East Sakhalin, Tarayka (Sisratoka)

Pilsudski renders *japaxcí* as ‘sailed’; we take the verb to be *yapa* ‘to move by boat, put in to shore’ (cf. Hokkaidō *yan/yap* ‘come ashore’), and *heyao* to contain the deictic *he-* ‘headed towards’ with *ya* ‘shore’.

The dictated *ucaskuma* differs sharply in texture from Asai Take’s *tuytah*. There is no reportative framing: where the west-coast oral tale stamps nearly every sentence with *manu*, Sisratoka’s narration runs in plain zero-marked third-person clauses, the evidential burden carried by the genre itself (Chapter 16) — a difference of elicitation as much as of dialect, since dictation suppresses the performative particles of live telling (Dal Corso 2024: 59). Aspect is articulated by the perfective auxiliary *hemaka* ‘finish’ (28) (Chapter 15); modality and negation by *koyakus* ‘be unable’ (30) and the volitional negative verb *etunne* ‘refuse, not want’ (27) (Chapter 17). The possessive forms *kotan-u* and *utar-hi*, with the *h* of the suffix elided, illustrate the morphophonology of the belonging form (Chapter 7, Chapter 5), and plural *-(a)hci* on *ray-ahci* and *yapa-hci* tracks the plural subjects that the zero third person leaves unexpressed (Chapter 8, Chapter 13). For the dialect position of Tarayka within East Sakhalin, and for the other tales of the collection, see Chapter 2.

26.4 The Dobrotvorsky dialogues (1875)

The Ainu–Russian dictionary of Mikhail Dobrotvorsky (Dobrotvorsky 1875), published at Kazan in 1875, carries in its appendix two short connected texts — the oldest connected prose in Sakhalin Ainu, predating Pilsudski’s fieldwork by a generation. The first is a scene of confession told by a man named Ciwokanke; the second a dialogue of request and marriage. Both come from the Kusunnay area of the west-central coast. They were re-edited, with romanisation, Japanese translation, and philological commentary, by Sakaguchi Ryō (Sakaguchi 2021), whose reading we follow throughout; the accented romanisation shown with the examples is his. We give four sentences from the first text and two from the second.

As with Sentoku’s letters (§26.2), the Cyrillic record is strong exactly where the grammatical interest lies. Dobrotvorsky’s transcription registers syllable-final consonants faithfully, but it ignores vowel length and wavers between *и* and *е* and between *о* and *у*, so that contrasts hinging on vowel quality or quantity cannot always be recovered from the spelling. The person markers — *an=*, *-an*, *in=*, *i=*, *e=* — live on consonants, and these the texts attest beyond doubt: for the morphology they are cited for here, the dialogues are more reliable witnesses than their age and script would suggest (Sakaguchi 2021).

- (32)
- Tan kotàn ohtà utása-an kusú áriki-an.*

tan kotan ohta utasa-an kusu ariki-an
 this village in visit-INDEF.S PURP come.PL-INDEF.S

‘I have come in order to visit this village.’ — Sakaguchi 2021: 49; West Sakhalin, Kusunnay (Ciwokanke)

The INDEF-series -an stands on both verbs, exactly as in Sentoku’s letters thirty years later; the purposive *kusu* links them. *ariki* is the plural stem of ‘come’: the indefinite person selects plural stems even under singular reference.

- (33)
- Tan kotán ta eh renkáyne, an-è-núkara.*

tan kotan ta eh renkayne an=e=nukara
 this village LOC come thanks.to INDEF.A=2SG.O=see

‘Thanks to your coming to this village, I have been able to see you.’ — Sakaguchi 2021: 49; West Sakhalin, Kusunnay (Ciwokanke)

an=e= stacks the INDEF agent on the second-person object, the transitive ‘I...you’ frame; *renkayne* ‘thanks to, because of’ is the same connective that runs through Sentoku’s letters and Sisratoka’s narrative above.

- (34)
- Tamà in-konte. Rékoro yay’iráyki-re-an.*

tama in=konte reekoro yay-irayki-re-an
 bead 1PL.EXCL.O=give very REFL-thank-CAUS-INDEF.S

‘You gave me a bead. I am very grateful.’ — Sakaguchi 2021: 49; West Sakhalin, Kusunnay (Ciwokanke)

The diagnostic line: *in=* ‘me’, the first-person object prefix of the east-coast paradigms, inside an otherwise INDEF-marked passage. *yay-irayki-re-an* ‘I cause myself to be thankful’ stacks reflexive, causative, and the INDEF -an; *Rékoro* for *reekoro* shows the script’s blindness to vowel length.

- (35)
- Pírikano omán ciki pírika.*

pirika-no oman ciki pírika
 good-ADV go.SG if good

‘It would be good if you went carefully.’ — Sakaguchi 2021: 49; West Sakhalin, Kusunnay (Ciwokanke)

A parting formula: *pirika* adverbialized with -no in the protasis and predicative in the apodosis, hinged on the conditional *ciki* — the same *ciki(n)* as in Sentoku’s opening request (15).

- (36)
- Anokay an-renkayne, i-koyki ánpe ka hánne.*

anokay an=renkayne i=koyki anpe ka hanne
 INDEF INDEF.A=accordingly INDEF.O=hit NMLZ even NEG

‘In accordance with my wish, I was not hit.’ — Sakaguchi 2021: 50; West Sakhalin, Kusunnay

The second text uses *i=* ‘me’ — the INDEF-series object, doubled by the free pronoun *anokay* — where the first text had *in=* (34); Sakaguchi notes that the dictionary itself records the parallel *kuan* ... *in-koyki*, a minimal pair between the two object prefixes. The negation is the analytic *hanne* over the nominalized *anpe ka* ‘even the fact of...’.

- (37)
- sine utah-ta ne kusu porono ham e=(i=)ye yahka pírika nankoro pe ne*
-
- one kin-? COP because much NEG 2SG.A=(INDEF.O=)say even.if good probably NMLZ COP

‘Since we are close kin, it would be fine even if you did not say much to me.’ — Sakaguchi 2021: 52; West Sakhalin, Kusunnay

sine utahta ne kusu is a fixed idiom that Dobrotvorskij himself enters in the dictionary as *sne utahta ne kusu* ‘because we are relatives’: *utah* is the concept form of *utar(a)* ‘people, kin’, so the literal sense is ‘being one set of kinfolk’ (Sakaguchi 2021, n. 43). The function of the frozen final -ta is unclear — plausibly the locative -ta in a lexicalized idiom — hence the ? in its gloss. The parenthesized *i=* on the morpheme line, and its (INDEF.O=) in the gloss, mark the editor’s restoration of an unwritten indefinite-object clitic. The negator here is *ham*, beside *hanne* in (36); the close *nankoro pe ne* ‘it will probably be (that...)’ is the epistemic formula.

Short as they are, the dialogues anchor the time-depth of everything else in this chapter. Their chief witness is the person system (Chapter 13): the INDEF-series runs throughout — -an on intransitives (32), the stacked *an=e=* ‘I...you’ frame (33) — and beside it stands *in=* ‘me’ (34), the object prefix elsewhere characteristic of the east coast (Satō 1985), so that this west-central material patterns with the East rather than with the north-west (Chapter 2). The negative system is already the analytic one: *hanne* (36) and *ham* (37) negate in 1875 exactly as in the twentieth-century corpus, a fixed point for the history traced in Chapter 17. The politeness formulae run on the clause-linking inventory — purposive and causal *kusu*, conditional *ciki*, concessive *yahka* (Chapter 23) — and the epistemic close *nankoro pe ne* stacks the modal *nankoro* on the nominalizer *pe* and the copula *ne*

(Chapter 15, Chapter 20, Chapter 18). Genre matters here too: where the folktales above are narrative, these are face-to-face exchanges, and their grammar is correspondingly that of address — second persons, softened requests, and not a single *manu*.

Appendices

Abbreviations and conventions

Interlinear glossing

Interlinear examples follow the Leipzig Glossing Rules. In the morphemic line, a hyphen - separates affixes from stems and an equals sign = marks clitic boundaries. In the gloss line, grammatical categories appear in small capitals and lexical glosses in lower case; a period joins several gloss elements that correspond to a single unsegmentable morpheme (as in *an* ‘exist.SG’). Every Ainu word and morpheme in an example is linked to the dictionary of this site, and every example carries its source and a dialect tag.

An asterisk * marks an ungrammatical form; a preceding question mark marks a form of doubtful acceptability. Examples labelled *constructed example* are illustrations assembled for expository purposes and are not attested utterances; all other examples are attested in the cited source.

Gloss abbreviations

1 first person	IMP imperative
1PL first person plural	INCL inclusive
1SG first person singular	INDEF indefinite person (the Sakhalin Ainu <i>an</i> / <i>-an</i> / <i>i-</i> set)
2 second person	INFR inferential evidential
2PL second person plural	INS instrumental
2SG second person singular	INT intensifier
3 third person	INTERJ interjection
3PL third person plural	IPFV imperfective
3SG third person singular	ITR iterative
A agent-like argument of a transitive verb	LOC locative
ADM admonitive (‘lest’)	MIR mirative
ADV adverbial(izer)	NEG negative
ANTIP antipassive	NMLZ nominalizer/nominalization
APPL applicative	O object (patient-like argument of a transitive verb)
ASS associative plural	ORD ordinal
ATTR attributive	PASS passive
CAUS causative	PERS personal (evidential of personal knowledge)
CL classifier (numeral counter suffix)	PFV perfective
COMIT comitative	PL plural
COMP complementizer	POL polite
COND conditional	POSS possessive (the “belonging” form of nouns)
CONJ conjunctive (clause-linking) particle	PRF perfect
COP copula	PROG progressive
DES desiderative	PROH prohibitive
DIM diminutive	PROX proximal
DIST distal	PURP purposive
EMPH emphatic	Q question particle
EP epenthetic segment	QUOT quotative
EVID evidential	RECP reciprocal
EXCL exclusive	REFL reflexive
FIN final (sentence-final) particle	REP reportative evidential
FOC focus	RES resultative
HORT hortative	S single argument of an intransitive verb
HUM human (in classifiers)	

SG singular
TOP topic
VBLZ verbalizer

Dialect tags

Tag	Meaning
West Sakhalin	Examples from West-coast varieties (Rayciska, Esutoru, Otasu, Odasu and neighbouring settlements).
East Sakhalin	Examples from East-coast varieties (Ay, Otasan, Tunayci, Ociho, Tarayka and neighbouring settlements).
Sakhalin	Examples whose locality is not recorded or which hold across Sakhalin varieties.

Other conventions

- Ainu forms are cited in the romanisation set out in the chapter on orthography; phonemic representations appear between slashes and phonetic ones in square brackets.
- Japanese and Russian names and titles are given in the original script at first mention, e.g. 樺太 *Karafuto*, Сахалин *Sakhalin*.
- Section references of the form §4.3 point within the present grammar; chapter titles are linked in the running text.

References

Entries marked reported are cited second-hand, through the source named in the citation, and were not consulted directly.

Sakhalin Ainu

Primary sources and studies of the language this grammar describes.

- Ainu Bunka Shinkō Kenkyū Suishin Kikō アイヌ文化振興・研究推進機構 (FRPAC). 2014. アイヌ語教材テキスト「カラフトのアイヌ語」(入門・初級・中級) [Ainu language teaching materials: Sakhalin Ainu (introductory, beginner and intermediate volumes)]. Sapporo: Foundation for Research and Promotion of Ainu Culture.
- Bugaeva, Anna. 2016. On the innovative nature of Sakhalin Ainu: Focusing on nominalization. *Studia Orientalia* 117: 95–117.
- Chiri Mashiho 知里真志保. 1942. アイヌ語法研究—樺太方言を中心として [A study of Ainu grammar, centred on the Sakhalin dialect]. 樺太庁博物館報告 [*Reports of the Karafuto Government Museum*] 4(4): 51–172. Toyohara. Reprinted in 知里真志保著作集 3 (Heibonsha, 1973).
- Dal Corso, Elia. 2018. Ainu evidentiality: A semantico-pragmatic analysis and a morphosyntactic account. *Doctoral dissertation, SOAS University of London*. London.
- Dal Corso, Elia. 2021. The development of analytic negatives in Sakhalin Ainu. In Jeon, Hae-Sung, Pete Sells, Zixi You, Sotaro Kita & Jaehoon Yeon (ed.), *Japanese/Korean Linguistics* 28. Stanford: CSLI Publications.
- Dal Corso, Elia. 2021a. The language and folklore of West Sakhalin Ainu: A re-edition of Kyōko Murasaki's *Karafuto Ainugo* with translation and grammatical notes. Munich: LINCOM. Annotated re-edition, with English translation, of Murasaki 1976.
- Dal Corso, Elia. 2024. Elements of Sakhalin Ainu phonetics, phonology, and morphosyntax in Bronisław Piłsudski's corpus of Ainu folklore. In Calvetti, Paolo (ed.), *Philological and linguistic analysis working together: Exploring language diversity in Japan (Ca' Foscari Japanese Studies 24)* 53–80. Venice: Edizioni Ca' Foscari.

- Dal Corso, Elia. 2025a. The mixed nominative-neutral-inverse morphological alignment of Sakhalin Ainu. *Studia Orientalia* 126: 187–218.
- Dal Corso, Elia. 2025b. Negation in Sakhalin Ainu: History and typology. *Studia Orientalia Electronica* 13: 38–64.
- Dobrotvorsky, Mikhail M. Добротворский, М. М.. 1875. Аинско-русский словарь[Ainu–Russian dictionary]. Kazan Казань: Kazan University.
- Furukawa Kyōko 古川恭子. 1967. アイヌ語カラフトライチシカ方言シンタクス[Syntax of the Sakhalin Raychishka dialect of Ainu]. 言語研究 [*Gengo Kenkyū: Journal of the Linguistic Society of Japan*] 50: 100–120. Furukawa Kyōko is the maiden name of Murasaki Kyōko.
- Hattori Shirō 服部四郎. 1961. アイヌ語カラフト方言の「人称接辞」について[On the “personal affixes” of the Sakhalin dialect of Ainu]. 言語研究 [*Gengo Kenkyū: Journal of the Linguistic Society of Japan*] 39: 1–20.
- Iwai Toshiaki 岩井俊昭, Asakura Toshimitsu 朝倉利光, Ifukube Tōru 伊福部達 & Kawashima Toshio 川嶋稔夫. 1987. ピウスツキ蠟管の音声再生と録音内容——レーザビーム反射法による光学式音声再生 [Sound-reproduction from Bronisław Piłsudski’s recordings and their contents: Reproduction of sound from old wax cylinders using the laser-beam reflection method]. 国立民族学博物館研究報告別冊 [*Bulletin of the National Museum of Ethnology, special issue*] 5: 147–163.
- Kitahara Jirōta Mokottunas 北原次郎太 (ed.). 2016. Wada Bunjirō: Karafuto ainu setsuwashū 3[Wada Bunjirō: Collection of Karafuto Ainu folktales, vol. 3]. Sapporo: Aynu Teetawanoankur Kanpinuye Cise. Cited after Dal Corso (2025b). reported
- Majewicz, Alfred F. (ed.). 1998. The collected works of Bronisław Piłsudski, vol. 3: Materials for the study of the Ainu language and folklore 2. *Trends in Linguistics: Documentation* 15-3. Berlin & New York: Mouton de Gruyter. Cited after Dal Corso (2025b). reported
- Murasaki Kyōko 村崎恭子. 1976. カラフトアイヌ語[Sakhalin Ainu]. Tokyo: Kokusho Kankōkai. Materials volume (資料篇), with cassette tapes.
- Murasaki Kyōko 村崎恭子. 1979. カラフトアイヌ語——文法篇[Sakhalin Ainu: Grammar volume]. Tokyo: Kokusho Kankōkai.
- Murasaki Kyōko 村崎恭子. 2009. 樺太アイヌ語の数詞について[On the numerals of Sakhalin Ainu]. In Tsumagari Toshirō 津曲敏郎 (ed.), サハリンの言語世界——北大文学研究科公開シンポジウム報告書 [*The linguistic world of Sakhalin*] 71–84. Sapporo: Graduate School of Letters, Hokkaido University.
- Murasaki Kyōko 村崎恭子. 2013. 藤山ハル口述 カラフトアイヌ語例文集 1[Short sentences in Sakhalin Ainu spoken by Fujiyama Haru, vol. 1]. Sapporo: Hokkaido University Center for Ainu and Indigenous Studies (online publication).
- Murasaki Kyōko 村崎恭子. 2025. エンチウ (樺太アイヌ) 語会話入門【改訂版】[An introduction to Enciw (Sakhalin Ainu) conversation, revised edition]. Kushiro: Fujita Insatsu Excellent Books.
- Murasaki Kyōko 村崎恭子 (ed. & tr.). 1996. 浅井タケ昔話全集 1：話者の絶えた樺太アイヌ語資料集成 [The complete folktales of Asai Take, vol. 1: A compilation of Sakhalin Ainu materials from a language whose speakers are gone]. Yokohama: 私家版 [privately published]. First edition; revised and completed as Murasaki 2001.
- Murasaki Kyōko 村崎恭子 (ed.). 2001. 改訂浅井タケ昔話全集 I・II[The complete folktales of Asai Take, vols. I–II, revised edition]. *Endangered Languages of the Pacific Rim (ELPR) Publication Series A2-010*. Suita: Osaka Gakuin University.
- Piłsudski, Bronisław. 1912. Materials for the study of the Ainu language and folklore. Cracow: Imperial Academy of Sciences.
- Piłsudski, Bronisław. 1998. The collected works of Bronisław Piłsudski, vol. 2: Materials for the study of the Ainu language and folklore (Cracow 1912). In Majewicz, Alfred F. (ed.), *Trends in Linguistics: Documentation* 15-2. Berlin & New York: Mouton de Gruyter.
- Sakaguchi Ryō 阪口諒. 2019. 山本多助筆録アイヌ語樺太方言テキスト(1)——「カラフト・ウベベケレ (オプケネワイコロウペベケレ)」[A Sakhalin Ainu text recorded by Yamamoto Tasuke (1): the Karafuto upepekere “Opke newa ikoro upepekere”]. 北方人文研究 [*Journal of the Center for Northern Humanities*] 12: 111–121. Sapporo: Hokkaido University.

- Sakaguchi Ryō 阪口諒. 2020a. アイヌ語樺太方言における utara の用法[The uses of utara in Sakhalin Ainu]. 北方言語研究 [*Hoppō gengo kenkyū: Northern language studies*] 10: 187–202.
- Sakaguchi Ryō 阪口諒. 2020b. アイヌ語樺太方言における名詞複数接尾辞 ahcin の用法[The uses of the nominal plural suffix -ahcin in Sakhalin Ainu]. 北方人文研究 [*Journal of the Center for Northern Humanities*] 13: 1–19.
- Sakaguchi Ryō 阪口諒. 2020c. ピウスツキ採集のアイヌ語樺太方言民話テキスト——「カレイ男とカジカ男」[A Sakhalin Ainu folktale text collected by Pilsudski: “The flatfish man and the sculpin man”]. 千葉大学大学院人文公共学府研究プロジェクト報告書 354 『ユーラシア諸言語の類型論的記述』 [*Chiba University research project reports 354: Typological description of Eurasian languages*] 43–69. Chiba.
- Sakaguchi Ryō 阪口諒. 2021. 『アイヌ語ロシア語辞典』中のアイヌ語樺太方言テキスト[Sakhalin Ainu texts in the Ainu–Russian dictionary (of Dobrotvorsky)]. アイヌ語・アイヌ文化研究の課題 [*Issues in Ainu language and culture studies*] (千葉大学大学院人文公共学府研究プロジェクト報告書 358) 43–55. Chiba: Chiba University.
- Sakaguchi Ryō 阪口諒. 2022. アイヌ語樺太方言における数詞と計算[Numerals and calculation in Sakhalin Ainu]. 北方人文研究 [*Journal of the Center for Northern Humanities*] 15: 63–84.
- Sakaguchi Ryō 阪口諒. 2024. 樺太アイヌ語における人称と数の標示に関する研究[A study of person and number marking in Sakhalin Ainu]. *Doctoral dissertation, Chiba University.*
- Satō, Tomomi. 1985. The first person objective affix in- in the East Coast dialects of Sakhalin Ainu. *Proceedings of the International Symposium on B. Pilsudski's Phonographic Records and the Ainu Culture* 157–167. Sapporo: Hokkaido University.
- Sentoku Tarōji 千徳太郎治. 1929. 樺太アイヌ叢話[Collected accounts of the Sakhalin Ainu]. Tokyo: Shikōdō Ichikawa Shōten 市光堂市川商店.
- Tamura Masashi 田村雅史. 1999. 樺太アイヌ語の借用語について——ウイльта語・ニヴフ語との関係を中心に[On loanwords in Sakhalin Ainu, with special reference to the Uilta and Nivkh connections]. 北方言語研究 [*Northern language studies*] 別冊. Cited after Shiraishi & Tangiku (2022). reported
- Tangiku Itsuji 丹菊逸治. 1998. アイヌ語樺太西海岸方言の-rV 音節で終る動詞について[On verbs ending in -rV syllables in West-coast Sakhalin Ainu]. 千葉大学ユーラシア言語文化論集 [*Chiba University Eurasian languages and cultures studies*] 1: 222–234.
- Tangiku Itsuji 丹菊逸治 & Ogihara Shinko 荻原眞子. 2001. 千徳太郎治のピウスツキ宛書簡——「ニシパ」へのキリル文字の手紙[Sentoku Tarōji's letters to Pilsudski: Letters in Cyrillic script to “nispa”]. 千葉大学ユーラシア言語文化論集 [*Chiba University Eurasian languages and cultures studies*] 4: 187–226.

Hokkaidō Ainu

Works on Hokkaidō varieties, cited for contrast only.

- Batchelor, John. 1938. An Ainu–English–Japanese dictionary. Tokyo: Iwanami Shoten. 4th edition.
- Bugaeva, Anna. 2012. Southern Hokkaido Ainu. In Tranter, Nicolas (ed.), *The languages of Japan and Korea* 461–509. London: Routledge.
- Dal Corso, Elia. 2020. The interaction of relativization and noun incorporation in Southern Hokkaidō Ainu. In Pappalardo, Giuseppe & Patrick Heinrich (ed.), *European approaches to Japanese language and linguistics*. Venice: Edizioni Ca' Foscari.
- Refsing, Kirsten. 1986. The Ainu language: The morphology and syntax of the Shizunai dialect. Aarhus: Aarhus University Press.
- Satō Tomomi 佐藤知己. 2008. アイヌ語文法の基礎[Foundations of Ainu grammar]. Tokyo: Daigaku Shorin.
- Tamura, Suzuko. 2000. The Ainu language. Tokyo: Sanseido.
- Urata Yū 浦田遊 (ed.). 1998. アイヌ・モシリ——幻のアイヌ語誌復刊[Ainu mosir: Reprint of the lost Ainu-language journal]. Kushiro: Kushiro Ainu Bunka Konwakai.

Pan-Ainu, comparative and methodological works

Dialectology, reconstruction, typology, and conventions.

- Asai Tōru 浅井亨. 1974. Classification of dialects: Cluster analysis of Ainu dialects. 北方文化研究 [*Bulletin of the Institute for the Study of North Eurasian Cultures*] 8: 45–136. Sapporo: Hokkaido University.
- Austerlitz, Robert. 1976. L'appellation du renne en ghiliak. *Études finno-ougriennes* 13: 25–28. Cited after Janhunen (2022). reported
- Bugaeva, Anna. 2010. Ainu applicatives in typological perspective. *Studies in Language* 34(4): 749–801. reported
- Bugaeva, Anna. 2015. An equivalent of the standard-of-comparison relativization in Ainu.
- Bugaeva, Anna. 2017. Noun-modifying clause constructions in Ainu. In Matsumoto, Yoshiko, Bernard Comrie & Peter Sells (ed.), *Noun-Modifying Clause Constructions in Languages of Eurasia: Rethinking theoretical and geographical boundaries (Typological Studies in Language 116)* 203–250. Amsterdam & Philadelphia: John Benjamins.
- Bugaeva, Anna. 2021. Appositive possession in Ainu and around the Pacific. *Linguistic Typology* 25(2): 263–305.
- Bugaeva, Anna (ed.). 2022. Handbook of the Ainu language. *Handbooks of Japanese Language and Linguistics* 12. Berlin & Boston: De Gruyter Mouton.
- Chiri Mashiho 知里真志保. 1952. アイヌ語に於ける母音調和[Vowel harmony in Ainu]. 北方文化研究報告 [*Reports on Northern Culture Studies*] 7. Sapporo: Hokkaido University. Page references follow the reprint in 知里真志保著作集 4 (Heibonsha, 1974).
- Chiri Mashiho 知里真志保. 1973. 知里真志保著作集 3：生活誌・民族学編[Collected works of Chiri Mashiho, vol. 3]. Tokyo: Heibonsha.
- Comrie, Bernard, Martin Haspelmath & Balthasar Bickel. 2015. The Leipzig Glossing Rules: Conventions for interlinear morpheme-by-morpheme glosses. Leipzig: Max Planck Institute for Evolutionary Anthropology.
- Croft, William. 1991. The evolution of negation. *Journal of Linguistics* 27(1): 1–27. reported
- Givón, Talmy. 1978. Negation in language. In Cole, Peter (ed.), *Syntax and Semantics, vol. 9: Pragmatics* 69–112. New York: Academic Press. reported
- Givón, Talmy. 1980. The binding hierarchy and the typology of complements. *Studies in Language* 4(3): 333–377. reported
- Gruzdeva, Ekaterina. 1996. The linguistic situation on Sakhalin Island. In Wurm, Stephen A., Peter Mühlhäusler & Darrell T. Tryon (ed.), *Atlas of languages of intercultural communication in the Pacific, Asia, and the Americas, vol. II: Texts* 1007–1012. Berlin & New York: Mouton de Gruyter.
- Hattori Shirō 服部四郎. 1967. アイヌ語の音韻構造とアクセント[The phonological structure and accent of Ainu]. 音声の研究 [*Onsei no Kenkyū: Study of Sounds*] 13: 207–223. Phonetic Society of Japan.
- Hattori Shirō 服部四郎 (ed.). 1964. アイヌ語方言辞典[An Ainu dialect dictionary]. Tokyo: Iwanami Shoten.
- Hattori Shirō 服部四郎 & Chiri Mashiho 知里真志保. 1960. アイヌ語諸方言の基礎語彙統計学的研究[A lexicostatistic study of the Ainu dialects]. 民族学研究 [*Japanese journal of ethnology*] 24(4): 307–342.
- Itabashi Yoshizō 板橋義三. 2001. 樺太アイヌ語の母音の長短と北海道アイヌ語のピッチアクセントとの史的関係 (1) —両アイヌ方言の音韻における史的関係の解明に向けての研究序説[The historical relation between Sakhalin Ainu vowel length and Hokkaidō Ainu pitch accent, part 1]. 言語文化論究 [*Studies in Languages and Cultures, Kyushu University*] 14: 87–103.
- Janhunen, Juha A.. 2022. Ainu ethnic origins. In Bugaeva, Anna (ed.), *Handbook of the Ainu language*. Berlin & Boston: De Gruyter Mouton.
- Kirikae Hideo 切替英雄. 1984. アイヌ語の名詞句の構造と合成名詞[The structure of noun phrases and compound nouns in Ainu]. 言語研究 [*Gengo Kenkyū: Journal of the Linguistic Society of Japan*] 86: 105–121.
- Kirikae Hideo 切替英雄. 1997. アイヌによるアイヌ語表記[The writing of Ainu by the Ainu]. 国文学 解釈と鑑賞 [*Japanese literature: interpretation and appreciation*] 62(1): 99–107. Tokyo: Shibundō.
- Kitahara Jirōta 北原次郎太, Tamura Masashi 田村雅史, Tamura Masato 田村将人, Tangiku Itsuji 丹菊逸治 & Tamura Suzuko 田村すず子 (eds.). 2003.

アイヌ語樺太・名寄・釧路方言の資料—藤山ハルさん・山田ハヨさん・北風磯吉さん・徹辺重次郎さんの口頭文芸・語彙・民族誌[Materials on the Sakhalin, Nayoro and Kushiro dialects of Ainu: Oral

- literature, vocabulary and ethnography of Fujiyama Haru, Yamada Hayo, Kitakaze Isokichi and Tetsube Jūjirō]. *ELPR Publication Series A2-039*. Suita: Osaka Gakuin University. Recordings by Tamura Suzuko.
- Lee, Sean & Toshikazu Hasegawa. 2013. Evolution of the Ainu language in space and time. *PLOS ONE* 8(4): e62243.
 - Moseley, Christopher (ed.). 2010. *Atlas of the world's languages in danger*, 3rd edn. Paris: UNESCO Publishing.
 - Nakagawa Hiroshi 中川裕. 2006. アイヌ人によるアイヌ語表記への取り組み[Ainu efforts at writing the Ainu language]. 表記の習慣のない言語の表記 [*Writing unwritten languages*]. Tokyo: Research Institute for Languages and Cultures of Asia and Africa (ILCAA), Tokyo University of Foreign Studies.
 - Nakagawa Hiroshi 中川裕. 2024. アイヌ語広文典[Ainugo kōbunten [An extensive grammar of the Ainu language]]. Tokyo: Hakusuisha 白水社.
 - Noonan, Michael. 2007. Complementation. In Shopen, Timothy (ed.), *Language typology and syntactic description*, vol. II: *Complex constructions (2nd edition)* 52–150. Cambridge: Cambridge University Press. reported
 - Ochiai Izumi 落合いずみ. 2021. Ainugo no sūshi saikō: nijisshinpō ni okeru kahō-sanpō kara jōhō-sanpō e no kirikae[Ainu numerals reconsidered: the switch from undercounting to overcounting in the vigesimal system].
 - Ono, Yohei. 2020. Reconsideration of “major division” of Ainu dialects: A statistical reanalysis of Asai (1974). 北方言語研究 [*Hoppō gengo kenkyū: Northern language studies*] 10: 231–254.
 - Satō Tomomi 佐藤知己. 2015. アイヌ語の合成語のアクセント規則とその例外について[On the accent rules of Ainu compounds and their exceptions]. アイヌ語研究の諸問題 [*Problems in Ainu language research*] 1–13.
 - Shibatani, Masayoshi. 1990. *The languages of Japan*. Cambridge: Cambridge University Press.
 - Shiraishi, Hidetoshi & Itsuji Tangiku. 2022. Language contact in the north. In Bugaeva, Anna (ed.), *Handbook of the Ainu language*. Berlin & Boston: De Gruyter Mouton.
 - Tangiku, Itsuji. 2022. Differences between Karafuto and Hokkaido Ainu dialects. In Bugaeva, Anna (ed.), *Handbook of the Ainu language*. Berlin & Boston: De Gruyter Mouton.
 - Veselinova, Ljuba. 2016. The Negative Existential Cycle viewed through the lens of comparative data. In van Gelderen, Elly (ed.), *Cyclical change continued (Linguistics Today 227)* 139–187. Amsterdam: John Benjamins. reported
 - Vovin, Alexander. 1993. *A reconstruction of Proto-Ainu*. Leiden: Brill.