

A Historical Dictionary of Yukaghir



Trends in Linguistics Documentation 25

Editors

Walter Bisang

Hans Henrich Hock

Werner Winter

Mouton de Gruyter
Berlin · New York

A Historical Dictionary of Yukaghir

by

Irina Nikolaeva

Mouton de Gruyter
Berlin · New York

Mouton de Gruyter (formerly Mouton, The Hague)
is a Division of Walter de Gruyter & Co., Berlin.

The publication of this volume was made possible by subsidies from
the *Finno-Ugric Society* and the *Foundation for Endangered Languages*.

⊗ Printed on acid-free paper which falls within the guidelines
of the ANSI to ensure permanence and durability.

Library of Congress Cataloging-in-Publication Data

Nikolaeva, I. A. (Irina Alekseevna)
A historical dictionary of Yukaghir / by Irina Nikolaeva.
p. cm. – (Trends in linguistics. Documentation ; 25)
Includes bibliographical references and index.
ISBN-13: 978-3-11-018689-5 (hardcover : alk. paper)
ISBN-10: 3-11-018689-6 (hardcover : alk. paper)
I. Yukaghir language – History – Dictionaries. I. Title. II. Series.
PM20.N55 2006
494'.6–dc22

2006012647

Bibliographic information published by Die Deutsche Bibliothek

Die Deutsche Bibliothek lists this publication in the Deutsche Nationalbibliografie;
detailed bibliographic data is available in the Internet at <<http://dnb.ddb.de>>.

ISBN-13: 978-3-11-018689-5

ISBN-10: 3-11-018689-6

ISSN: 0179-8251

© Copyright 2006 by Walter de Gruyter GmbH & Co. KG, D-10785 Berlin

All rights reserved, including those of translation into foreign languages. No part of this
book may be reproduced or transmitted in any form or by any means, electronic or mechan-
ical, including photocopy, recording or any information storage and retrieval system, with-
out permission in writing from the publisher.

Printed in Germany.

Acknowledgements

I would like to express my gratitude to all my Yukaghir consultants, and especially my principle guide to the language, the late Vasilij Salugin. I am deeply indebted to Tapani Salminen for his help in converting my database into a format suitable for publication. Without his assistance and encouragement this volume would not have appeared. I also thank Aleksandr Anikin, Uwe Bläsing, Michael Fortescue, Eugene Helinski, and Tapani Salminen for discussing Yukaghir materials and related questions with me. Many special thanks go to Simon Carne for patiently editing the English text. The work on the book has been partly supported by the Spinoza Prize awarded to Frederick Kortlandt by the Netherlands Organization for Scientific Research, by the Finno-Ugric Society (Helsinki), and by Deutsche Forschungsgemeinschaft through Sonderforschungsbereich 471 at the University of Konstanz. I wish to thank Professor Kortland for his kind support of my work. I am very grateful to the Foundation of Endangered Languages for sponsoring this publication. The book is dedicated to the memory of my father, Aleksej Vasil'evič Nikolaev.

Preface

Yukaghir is spoken in the extreme North-East of Siberia. Several centuries ago there used to be several Yukaghir idioms, but by the end of the 19th century only two varieties had survived. They are referred to as Tundra (or Northern) Yukaghir and Kolyma (or Southern) Yukaghir. These languages exhibit fairly significant differences, especially in vocabulary. Although traditionally they have been termed dialects, some scholars prefer to view them as separate languages and speak about “a Yukaghir language family” rather than “the Yukaghir language” (Kurilov 2001; Maslova 2003). According to estimates from the late 1980s, Tundra Yukaghir is spoken fluently by about 150 people and Kolyma Yukaghir by about 50 people, but these numbers seem to be rapidly decreasing.

This book has two main purposes. First, it is intended as a relatively complete source of information on the Yukaghir lexicon. No comprehensive lexicographic description of Yukaghir is available for the international linguistic community. In addition to several relatively small glossaries (Veenker 1989; Endo 1997, 2001; Nyikolajeva 2000; Nikolaeva and Šalugin 2003; Maslova 2001, 2003), two major dictionaries have been published so far: Angere (1957) and Kurilov (2001). The former is based on Jochelson (1900) and mostly presents the data from Kolyma Yukaghir as spoken at the end of the 19th century. The transcription and morphological segmentation used in this dictionary are seriously flawed. Kurilov (2001), together with its shorter version Kurilov (1990), is a very rich collection of the modern Tundra Yukaghir lexicon, but it uses the Cyrillic transcription with Russian translations and copies are rare outside Russia. Most importantly, each of these sources only describes one Yukaghir idiom. In contrast, this dictionary presents the different varieties of Yukaghir in comparative format. The modern Tundra Yukaghir materials are taken from published sources, while the modern Kolyma data were obtained through my own fieldwork conducted in the 1980s–1990s. Although some lacunae are inevitable, the book presents the first fairly comprehensive lexicographic description of Kolyma Yukaghir. In this respect it can be viewed as an attempt to document the lexicon of this highly endangered language. The dictionary also contains earlier materials starting from the late 17th century, most of which reflect extinct varieties of Yukaghir.

Second, and most importantly, the dictionary provides a reconstruction of Proto-Yukaghir, which has not previously been attempted (but see Nikolaeva 1988). Proto-Yukaghir is understood as the common ancestor of all known Yukaghir varieties. Proto-Yukaghir stems are established based on internal reconstruction and comparison between various Yukaghir idioms, as well as the external data. In some cases I cite possible cognates of Yukaghir stems from other languages, mainly Uralic and Altaic. Yet this dictionary is not etymological in the usual sense, i.e. it is not meant to provide etymologies for all Yukaghir words. Our present knowledge of the history of Yukaghir is insufficient for this purpose, but this dictionary may be a first step in this direction.

The Yukaghir family is usually classified as isolated, although its possible distant genetic relationship with Uralic has been much discussed (Collinder 1940, 1957, 1965; Bouda 1940; Angere 1956; Tailleur 1959b; Nikolaeva 1988; Fortescue 1998). There is no consensus on this matter, mainly because the history of Yukaghir has remained little known. Since in this volume Yukaghir forms are not only cited in their modern shape, but are reconstructed for the first time, my hope is that it will provide a foundation for future etymological work. The dictionary will assist scholars to establish the affiliation of Yukaghir and to reconstruct the ethnic and linguistic situation in prehistoric northern Asia, as well as to investigate the possibility of distant genetic relations between language families in general.

Table of contents

Acknowledgements	v
Preface	vii
Abbreviations	xi
Introduction	1
1. Organization of the dictionary	1
1.1. The structure of an entry	1
1.2. Sources	2
1.3. Presentation of the data and translations	3
1.4. Transcription and transliteration	5
1.5. Grammatical information	7
1.6. Reconstructions, etymologies, and comments	8
2. Sources of the Yukaghir material	10
2.1. Old Yukaghir	10
2.1.1. Witsen	10
2.1.2. Mueller/Lindenau	12
2.1.3. Klitschka	15
2.1.4. Boensing	17
2.1.5. Billings/Merk	18
2.1.6. Matjuskin (Čuvan and Omok)	24
2.2. End of the 19th century and 20th century	25
2.2.1. Kurilov	25
2.2.2. Krejnović	25
2.2.3. Spiridonov	25
2.2.4. Jochelson	26
2.2.5. Schiefner	27
3. Basics of Kolyma Yukaghir phonology	29
3.1. Phonemic inventory	29
3.1.1. Vocalism	29
3.1.2. Consonantism	31
3.1.3. Some non-phonemic variations	33
3.2. Positional restrictions on consonants	34
3.3. Vowel harmony	35
3.3.1. Harmony in the first foot	36
3.3.2. Harmony of short vowels after the first foot	37

3.3.3. Harmony of long vowels	39
3.3.4. Synharmonism	40
3.4. Alternations of vowels	41
3.4.1 Quantitative alternations	41
3.4.2. Deletion	43
3.4.3. Assimilation of mid vowels	45
3.4.4. Reduction of the stem-final vowel	46
3.5. Alternations of consonants	47
3.5.1. “Sonorant ~ obstruent” alternation	47
3.5.2. Assimilation of voiced consonants	48
3.5.3. Voicing and palatal assimilation	50
3.5.4. Assimilation of sibilants and <i>j</i>	51
3.6. Epenthesis	52
3.6.1. Consonantal epenthesis	53
3.6.2. Vocalic epenthesis in word edge clusters	53
3.6.3. Word-medial vocalic epenthesis	54
4. A reconstruction of Proto-Yukaghir phonology	57
4.1. Vocalism	57
4.1.1. Inventory and vowel harmony	57
4.1.2. Non-high vowels	57
4.1.3. High vowels	61
4.1.4. Long vowels	64
4.2. Consonantism	65
4.2.1. Inventory	65
4.2.2. Sibilants and affricates	66
4.2.3. Fricatives	68
4.2.4. Sonorants and glides	69
4.2.5. Consonantal clusters	71
4.3. The structure of the stems	74
4.3.1. Monosyllabic stems	74
4.3.2. Bisyllabic stems	75
4.3.3. Trisyllabic stems	77
5. List of affixes	79
Notes	84
References	87
Dictionary	95
Index of meanings	465
Language index	499

Abbreviations

Languages

Chuk.	Chukchi	FP	Finno-Permic
Esk.	Eskimo	FU	Finno-Ugric
Ev.	Even	FV	Finno-Volgaic
Evk.	Evenki	MG	Proto-Mongolian
K	Kolyma Yukaghir	NS	Northern Samoyed
Kor.	Koryak	NT	Northern Tungus
Mong.	Modern Mongolian	OU	Ob-Ugric
Nen.	Nenets	OY	Old Yukaghir
Ngan.	Nganasan	PA	Altaic
Ost.	Ostjak	PE	Proto-Eskimo
Rus.	Russian	PY	Proto-Yukaghir
Selk.	Selkup	S	Samoyed
T	Tundra Yukaghir	TU	Tungus
Yak.	Yakut	U	Uralic

Sources of Yukaghir materials

B	Billings (1787)
BO	Boensing (1781)
KD	Kolyma Yukaghir from Jochelson's manuscript dictionary
KJ	Kolyma Yukaghir (Jochelson 1898, 1900)
KK	Kolyma Yukaghir (Krejnović 1982)
KL	Klitschka (1781)
M	Maydell (Schiefner 1871b)
MC	Chuvan materials of Matjuskin (Wrangel 1841)
ME	Merk (1787)
MK	Kolyma materials of Mueller/Lindenau (1741)
MU	Ust'-Janskoe materials of Mueller/Lindenau (1741)
MO	Omok materials of Matjuskin (Wrangel 1841)
RS	Rajskij/Stubendorf (Schiefner 1871a)
SD	Kolyma Yukaghir (Spiridonov 2003)
SU	Suvorov (Schiefner 1871a)
TD	Tundra Yukaghir from Jochelson' manuscript dictionary

TJ	Tundra Yukaghir (Jochelson 1900)
TK	Tundra Yukaghir (Krejnović 1958, 1982)
W	Witsen (1692)

Sources of etymologies

CED	Fortescue et al. (1994)	JRS	Slepcev (1972)
DEWOS	Steinitz (1966–1993)	JU	Collinder (1940)
EDAL	Starostin et al. (2003)	HUV	Collinder (1965)
ESRD	Anikin (2000)	LR	Fortescue (1998)
ESRZ	Anikin (2003)	UJN	Collinder (1957)

Grammatical terms

ABL	ablative	ITER	iterative
ACC	accusative	HAB	habitual
AJD	adjectival derivation	HORT	hortative
ADV	adverbial	LAT	lative
AN	action nominal	LOC	locative
APPL	applicative	MULT	multiplicative
ATTR	attributive	N	nominal derivation
AUGM	augmentative	NEG	negation
BP	body part	NOM	nominative
CAUS	causative	NONIT	non-iterative
COLL	collective	OF	object focus
COM	comitative	ORD	ordinal
COND	conditional	PART	participle
CONN	connective	PERF	perfective
CONV	converb	PL	plural
DAT	dative	POS	possessive
DEL	delimitative	PP	postposition
DES	desiderative	PRED	predicative
DIM	diminutive	PROH	prohibitive
DIR	directive	PROL	prolative
DS	different-subject converb	PRON	pronominal
EMPH	emphatic	QUAL	qualitative
EV	evidential	RES	resultative
FREQ	frequentative	SF	subject focus
FUT	future	SG	singular

GEN	genitive	SN	subject nominal
IMP	imperative	SS	same-subject converb
IMPF	imperfective	STAT	stative
INCH	inchoative	SUP	supine
INDEF	indefinite	TEMP	temporal
INST	instrumental	TR	transitive
INT	intensive	TRANS	translative
INTER	interrogative	INTJ	interjection
INTR	intransitive	V	verbal derivation

Other abbreviations

dial.	dialectal	rect.	rectus
id.	idem	rev.	reverse
lit.	literally	smb	somebody
p.	page	smth	something
pr.	proper noun		

Introduction

1. Organization of the dictionary

The dictionary part of the volume is preceded by this Introduction which explains the principles of the phonological transcription adopted in the book, describes the sources of the data, and discusses the grounds for the suggested reconstruction, as well as the relationship between different varieties of Yukaghir. The dictionary is followed by an English index intended to facilitate the search for Yukaghir equivalents of English words, and a language index.

1.1. The structure of an entry

The data in the dictionary are organized alphabetically according to reconstructed stems preceded by an asterisk. In the alphabetical order the diacritics such as the palatalization sign ('), the length sign (:) and the hachek are disregarded. The letters *o* and *o*; *n*, *n* and *n*; *g* and *g*; *l* and *l*; *d*, *d* and *d*, as well as *e* and *e* are treated as one letter each. For recent borrowings into Yukaghir, mainly from Russian, Yakut and Even, the source forms are cited without an asterisk. Russian verbs are cited in the infinitival form, although Yukaghir normally borrows the bare stem. Each reconstruction or source word is provided with a consecutive number from 1 to 2659, with which the entry is cited in the reverse index. Unfortunately, for technical reasons the following 36 numbers are missing from the numeration: 113, 196, 207, 318, 416, 428, 437, 588, 690, 799, 1114, 1143, 1144, 1222, 1244, 1296, 1559, 1578, 1706, 1736, 1873, 1922, 1923, 2001, 2080, 2119, 2175, 2216, 2259, 2266, 2460, 2461, 2462, 2463, 2494, 2503. Thus, the dictionary includes 2623 entries.

In each entry the reconstruction or the source word is followed by the Yukaghir data. The source of the data is indicated by an abbreviation in capital letters before the word. For most entries the data are organized in several groups, each of which is written on a separate line. A group is formed by at least two words from different sources which exhibit more or less identical morphological structure in an uninflected form. For example, verbs have the same aspectual affixes, and nouns have the same derivational affixes or are parts of the same compound. When deciding whether words belong to the

2 Introduction

same group I have ignored transcriptional differences between the sources, as well as regular phonological variations. The data within a group is cited in the following order: data from the 20th century Kolyma Yukaghir – data from the 20th century Tundra Yukaghir – data from the earlier sources.

The words in a group normally have the same meaning and it is only cited once, after the first word. So when the following words in a group are not accompanied by a translation, it should be assumed that they have the same meaning as the first word in this group. Where the meaning of a word in a group differs, a translation immediately follows. If a word has an additional meaning together with the meaning it shares with other words, this is indicated with a plus sign (+). Words from different sources within a group are separated by a semi-colon. As an example I cite below part of the entry for the stem **köŋ*-.

(1) **köŋ*-

K *köŋe*:- to chop; KK *köŋie*-, *koŋie*-; KJ *koŋie*-; KD *koŋie*-; T *köŋie*- to undo, to unrip; to cut; TJ *kaŋie*-; TD *konie*- to cut a skin or a fish

K *köŋu*: split, crack; KK *kuŋu*: + scratch; KD *koŋu*:

KJ *koŋda*- to take off | TK *köŋiere*- to cut

This fragment includes two groups, one with the basic meaning ‘to chop’ and another ‘split, crack’. Both go back to the reconstructed proto-stem **köŋ*-. Some words have meanings that differ from the basic meaning, e.g. TD *konie*- ‘to cut a skin or a fish’. KK *kuŋu*: means ‘split, crack’ and additionally ‘scratch’, as indicated by the plus sign. The entry contains also the words KJ *koŋda*- ‘to take off’ and TK *köŋiere*- ‘to cut’ derived from the same stem but having no exact correspondence in any other available sources. Such words are not members of a group. They are presented on a separate line and divided by the sign |. Translations are given after each word.

When it is unclear whether the word belongs to the relevant entry, it is preceded by a question mark. Some entries also contain brief etymological remarks and comments (see 1.6 for details). These are written on separate lines.

1.2. Sources

The modern Kolyma data come from my field materials recorded among the Kolyma Yukaghirs in the settlements Nelemnoe and Zyrjanka, as well as in Jakutsk, in 1986–1991. Some of these materials remain unpublished, others

have been published as text collections (Nikolaeva 1989, 1997, 2000) and a school dictionary (Nikolaeva and Šalugin 2003). I have also used other sources of modern Kolyma Yukaghir for consultation, such as Spiridonov (1997), Endo (1997, 2001), and Maslova (2001, 2003). However, I do not indicate them because they mostly correspond to my field data.¹

It is not my goal to present all Kolyma Yukaghir derivations here, especially the numerous aspectual forms of verbs, which exhibit various degrees of productivity. Normally, verbal derivations are cited in three cases: (i) if the derived form has a direct correspondence in at least one other source and therefore is part of a group, (ii) if its meaning is unpredictable and substantially differs from the base form, and (iii) if it provides some phonological information relevant for the reconstruction. So if a verbal derivational form is not cited, the reader should not automatically assume it to be non-existent. Relevant generalizations and a list of attested verbal derivations can be found in Maslova (2003). On the other hand, I have tried to cite all or most attested nominal derivatives and compounds, as well as some idiomatic expressions. Words of Russian origin in Kolyma Yukaghir are only cited if they are recorded several times from different informants and so can be considered more or less established loanwords. But I cite all Russian words that occur in other sources.

The modern Tundra Yukaghir data are exclusively taken from the works of Kurilov (1990, 2001). Other data from the 20th century and the late 19th century come from the works of Jochelson, Spiridonov and Krejnovič. My intention was to cover these sources as fully as possible, but I have omitted many expressions whose status seems to be in between lexicalized compounds and free collocations, as well as some predictable derivational forms.

Earlier materials are represented by the data from Schieffner (1859, 1871a, 1871b) from the second half of the 19th century, and by what I refer to as Old Yukaghir, i.e. data from all known Yukaghir sources starting from the late 17th century until the first third of the 19th century (on Old Yukaghir see 2.1). They are taken either from published sources or unpublished archive collections. See Abbreviations for explanations of the labels, and Section 2 of the Introduction for descriptions of the sources.

1.3. Presentation of the data and translations

Although Kolyma Yukaghir data are unified on phonological grounds, in some cases I also cite free and idiolectal variants; they are separated with a comma. Verbs are cited as uninflected stems followed by a hyphen, as are

some postpositional stems that can host an inflection. Nouns are given in the basic Nominative Singular form.

The words from published and unpublished sources are normally cited as they appear in the source, while Cyrillic-based data are transliterated (see 1.4). In some cases I have chosen to provide corrected versions of records that are obviously erroneous in square brackets after the abbreviation *rect.* Variations that occur in the sources are separated by commas.

The presentation of the data from the 20th century sources basically follows the same format as for modern Kolyma Yukaghir. Nouns are written without a hyphen, unless they only occur in an inflected form in the source in question. Verbal stems are given without inflectional affixes and are followed by a hyphen. Because of fusional processes on the morpheme boundaries, such truncated forms may differ from the actual stem. For example, KJ ‘strange, odd’ is represented by the participle *pailice*, where *-ce* is a participial affix. In fact the verbal stem is *pajlu:-*, while *-u:* changes into *-i* before certain affixes. But as the stem *pajlu:-* is not attested in KJ in this shape, I cite it as *pajli-*, merely omitting the participial suffix.

In contrast, pre-20th century data are given as they occur in the source, i.e. in inflected forms. For example, the stem *mon-* ‘to say’ is represented in SU as *mo'ny* and *monka*. These are inflected forms of the 3rd person Singular and Plural, respectively, cf. K *mon-i* and *mon-ŋi*. Translations and grammatical comments for such forms are not provided.

Hyphenation, spacing and capitalization that occurs in the sources is not preserved. However, I provide a morphological analysis for compounds and separate their components with a hyphen. Hyphenation concerns both nominal and verbal compounds. By verbal compounds I mean lexicalized expressions which normally include a verb and a noun in a possessive, case or postpositional form. Proclitics are also separated with a hyphen, while suffixes are not. In some cases it is difficult to see whether we are dealing with a compound or a suffix, e.g. *-ed'ilil* in T *aŋded'ilil* ‘waist’ (cf. K *aŋd-il* ‘waist’). Such elements are not separated with a hyphen.

If a compound is attested in more than one source, it is cited as a group of words in the entry that corresponds to its first component. Consider for example a fragment of the entry **aŋa*.

- (2) K *aŋan-pugelbe*: beard, moustache [lit. mouth fur]; KJ *aŋade-pugelbie*, *aŋa-pugolbie*; KD *aŋa(n)-pugelbie*, *aŋan-pudeli*; T *aŋa-buguče*; TK *aŋan-bugut'e*; TD *an'an-bugoče*, *an'an-pugoče*; B *anghen-bugu'lbr*; ME *angen-bugulbie*

As can be seen in (2), literal translations may be provided to facilitate the understanding of the morphological structure of a compound. Note that in most nominal compounds the first component stands in the Genitive in *-n/-d*, which does not belong to the stem.

If the second component of a compound does not exist independently, it is cited as a separate entry. If a compound is represented in only one source, for modern Kolyma and Tundra Yukaghir it is given as a whole. Isolated compounds from other sources are not cited, but their components are treated separately under their respective entries. Consider the following example.

(3) T *morqile* broken country with many hills and holes; TD *-morxile*

The TD *-morxile* is only attested as the second component of an isolated compound, i.e. a compound which does not have a direct correspondence in other sources. This compound is not given, but TD *-morxile* is cited together with the independently existing T *morqile* and is preceded by a hyphen.

Where the translations provided in the sources were in a language other than English, I have given English equivalents. However, I shortened and simplified some of the translations from Kurilov (2001), which gives several meanings for each entry. For fuller information on the semantics of Tundra Yukaghir words the reader should consult Kurilov's original works. Closely related meanings are separated by commas, while more distantly related meanings are separated by semi-colons. Obviously, the decisions I had to make here are somewhat individual. Proper nouns are translated with the abbreviation *pr.* and further explanations are given in brackets (e.g. a man, woman, river, etc). For most birds and fish Latin names are provided in addition to English translations.

1.4. Transcription and transliteration

The modern Kolyma Yukaghir data are cited in phonological transcription, as explained in Section 3 of this Introduction. Latin-based sources are TD, KD, B, ME, M, R, SU, ST, W, MU, and MC. These data are cited essentially as in the source, but I transliterated *ĩ* as *y*, denoted vowel length with a colon rather than a macron, and have ignored some diacritics (for example, the signs *^*, *~* and *°* on vowels). Some small changes in the citation of individual sources are also possible. They are explained in the respective subsections of Section 2.

Sources based on Cyrillic are T, KJ, TJ, TK, KK, SD, BO, KL, MC, and MO. The data from these sources are given in transliteration. In (4) I present the basic rules of transliteration applicable to all sources cited in this volume.

(4)	Cyrillic	Transliteration	Cyrillic	Transliteration
	а	a	п	p
	б	b	р	r
	в	w, v	с	s
	г	g	т	t
	с	γ	у	u
	д	d	ф	f
	е	e, je-	х	q, x
	ё	jo, 'o	ц	c
	ж	ž	ч	č
	з	z	ш	š
	и	i	щ	čš
	й	j	ъ	Ø
	к	k	ы	y
	л	l	ь	'
	м	m	э	e
	н	n	ю	ju, 'u
	ң	ŋ	я	ja, 'a
	о	o	е, “jat”	e, je-
	ө	ö		

This system differs slightly from the standard transliteration used to render Cyrillic in that *в* is rendered as *w* rather than *v* and *х* is rendered as *q* rather than *x*. Such a transliteration brings the records of the Yukaghir words closer to phonemic transcription. Note that when transliterating *в* and *х* in Russian words I use the letters *v* (rather than *w*) and *x* (rather than *q*), respectively.

The so-called jotacized letters (*ё*, *ю* and *я*) are transliterated in two ways depending on their position. Word-initially or after a vowel they are transliterated as the combination of *j* and the corresponding non-jotacized vowel (*o*, *u* or *a*). After a consonant they are transliterated as a non-jotacized vowel, while the preceding consonant takes a palatalization sign ('). Again the purpose of this system is to make the form closer to the Yukaghir equivalent. The jotacized vowel *e*, as well as the Old Russian letters *е* and “jat”, occur only in Old Yukaghir materials, since modern Cyrillic-based sources use *э*

instead. These letters do not normally involve palatalization of the preceding consonant. However, in Matjuškin's materials (MO and MC) the initial *e-* may denote *je-* as well as *e*, cf. MC *eme* ~ K *eče*: 'father' but MC *eōōũ* ~ T *jabo:j* 'dead'. For these sources I write *je-* when it is confirmed from other sources and *(j)e-* when the word is otherwise unknown.

The word-final "hard sign" *ъ* which was used after a final consonant in the old Russian orthography does not denote any sound and has been left out in the transliteration. The "soft sign" *ь* denotes palatalization of the preceding consonant, but if it occurs after *ч* (= *č*) palatalization is not marked because *č* is already palatalized. When this sign is used after *н*, *л*, and *д*, the whole string is transliterated as *ń*, *ĺ*, and *d'*, respectively. Boensing (BO) and Klitschka (KL) additionally use the sign *́*, whose meaning is not entirely clear. On vowels it is likely to denote stress, while on consonants palatalization. This sign is preserved in transliteration, but at the beginning of the word before a vowel it is rendered as *j*. Some Cyrillic letters, e.g. *ѣ*, *ѱ* and *ѱ*, occur in Yukaghir words only by mistake, since the corresponding sounds do not exist in Yukaghir.

1.5. Grammatical information

The dictionary is not intended as an introduction to Yukaghir grammar and provides only a minimum of grammatical information, mostly indications of word classes. Yukaghir distinguishes the following grammatical classes: (i) nouns, (ii) verbs, (iii) adjectives, (iv) adverbs, (v) postpositions, (vi) numerals, (vii) pronouns, (viii) conjunctions, (ix) particles, and (x) interjections.

Verbs and nouns are easily distinguished by morphological criteria. As mentioned above, nouns are uninflected in the Nominative Singular and are cited in this form. Verbs rarely appear uninflected, and verbal stems are followed by a hyphen. The category of verbs includes qualitative and quantificative stems translated with adjectives and numerals, respectively, e.g. K *embə-* 'black' and *ataql-* 'two'. Transitive and intransitive verbs have different inflectional paradigms. In some cases I indicate whether the Yukaghir verb is transitive or intransitive by the abbreviations TR and INTR after the translation. This is done, first, if the English translation is ambiguous, as e.g. for the verb 'to open', and second, when the transitivity of the Yukaghir verb does not correspond to its English equivalent. For example, K *ege:rə-* 'to tread on, to step on (TR)' is transitive, unlike its English counterpart.

Adjectives in Yukaghir form a small closed class, which includes at least the following words: K *ñančə* ‘big’, *juku* ‘small’, *čomo* ‘big’, *joloqə* ‘last’, *poñqə* ‘white’, *pömkə* ‘round’, *čickə* ‘long, tall’, *cö:lə* ‘old, ancient’, and *illə* ‘new, fresh, (an)other’. The grammatical class of these words is not indicated in the dictionary. Adverbs, numerals, conjunctions and pronouns are not indicated either, but their category should normally be clear from the translation provided. Postpositions in Kolyma Yukaghir are marked as PP. Interjections are marked as INTJ and in some cases an approximate translation or an explanation of use is provided in brackets.

Particles either express a certain grammatical meaning or serve discourse purposes. Unlike adverbs or interjections they cannot form an independent utterance, not even an elliptical one. From a morphological viewpoint they rather form an heterogeneous class. Some particles are free standing words, others are bound words, while yet others are clitics. I do not provide details of their morphological behavior here. Some particles are translated with their English equivalents, while for others I provide a description using the general word “marker”. For example, the approximate meaning of the K *qata* is something like ‘let us’, but in the dictionary it is translated as ‘hortative marker’. Proper nouns and interjections, as well as most particles, are not included in the reverse index.

Section 5 of this Introduction contains a list of Yukaghir inflectional and derivational morphemes which is meant to facilitate the morphological analysis of the cited materials. Grammatical labels used in the dictionary are largely based on Maslova (2003) with a few deviations discussed in Nikolaeva (2005).

1.6. Reconstructions, etymologies, and comments

The principles of the Proto-Yukaghir reconstruction are explained in Section 4 of this Introduction, which should be consulted before using the dictionary. Alternative reconstructions are indicated with a slash or brackets. Homonymous reconstructions are provided with numbers, e.g. **aj-* 1 and **aj-* 2. As with all sources, if a reconstructed stem does not normally occur in an uninflected form, it is followed by a hyphen. This is characteristic of all verbal stems, as well as some non-verbal stems. No attempt is made to reconstruct the Proto-Yukaghir meaning, but in most cases it can be easily deduced from the meanings of the daughter words.

If the word is only attested in one Old Yukaghir source, the reconstruction may only be a very rough approximation, since these sources

are not very reliable. Such reconstructions are marked with a question mark. Thus, the meaning of the question mark differs: in the presentation of the data it indicates that it is questionable whether the word is related to the relevant stem, while in a reconstruction it indicates that it is approximate.

In most cases I have tried to reconstruct a stem alone and so omitted easily recognizable and fairly productive affixes even if they are present in all daughter words. For example, the aspectual and valence-changing verbal affixes are excluded from the reconstruction. On the other hand, reconstructions may include some non-productive affixes, such as for example the nominal suffixes *-r/-t*, *-mə*, *-lə*, and *-kə/-qə*.

Etymological comments are very brief. I mostly cite potential cognates from Uralic and Tungus-Manchu languages, but in a few cases I also mention Turkic, Mongolian, Chukchi and Eskimo parallels. For the most part potentially related words are cited as they appear in the cited sources. In some instances small transcriptional changes are made, but without any substantial revisions. If the comparison with Yukaghir is highly problematic, it is preceded by a question mark. If a comparison has been suggested in the previous literature, references are given after the double slash //. When such references are missing, the comparison is suggested here for the first time. Translations are provided for the cognate words, unless their meanings are basically identical to the meaning of the Yukaghir words and are easily detected.

In the comment line I indicate whether the phonological shape of the word is irregular and comment on individual phonological and semantic changes. Irregularity either suggests that the word looks like a recent borrowing but the source is unknown to me, or that the word is morphologically complex. The comment line may also contain information on whether the Yukaghir word was borrowed into another language.

2. Sources of the Yukaghir material

This section of the Introduction describes the sources from which the Yukaghir data are cited, other than my own field notes on Kolyma Yukaghir, which will be addressed in Section 3.

2.1. Old Yukaghir

In this section I outline the sources of what I refer to as Old Yukaghir. This term describes the varieties of Yukaghir spoken between the first known record of the language (in the second half of the 17th century) and the first third of the 19th century. The upper “border” is marked by the time when an important phonological change, the simplification of the intervocalic clusters, took place (see 4.2.5).

The Old Yukaghir data come from published and unpublished sources. The unpublished sources are kept in the following archives: the archive of the Academy of Sciences, St Petersburg, abbreviated as AAN, the Manuscripts Department of the State Public Library, St Petersburg, abbreviated as OR GPB, and the Russian State Archive of Ancient Acts, Moscow, abbreviated as RGADA. Most of these data were collected by the expeditions of the Russian Academy of Sciences, such as the Second Kamchatkan Expedition (1733–1743), the North-Eastern Expedition of Joseph Billings (1785–1793), and the expedition of Ferdinand von Wrangel (1821–1824). The work of these expeditions is addressed in detail in Gnučeva (1940), Vdovin (1954) and Širina (1983), among others. Other materials were collected by Petr Pallas via correspondence with the local administration and later systematized by Friedrich Adelung, as part of the Russian Academy of Science’s programme for the compilation of comparative dictionaries of world languages. Pallas published his materials twice, in slightly different versions (Pallas 1786, 1789), while Adelung’s compilations remain in the archives.¹ All these sources represent slightly different idioms, but I leave open the question of whether they should be classified as dialects or separate languages. The philological analysis of Old Yukaghir sources presented here was first suggested in my dissertation (Nikolaeva 1988a).

2.1.1. *Witsen*

The first known record of Yukaghir is actually a published text. It appeared in a book by the Dutch diplomat and scholar Nicolaas Witsen (1641–1717) first

published in 1692. Witsen himself did not visit the Yukaghir land. He received the Yukaghir materials by post sometime after his visit to Moscow in 1664–1665, probably through his cousin Andrea Vinius, who later headed the Siberian Chancellery in Moscow (Gebhard 1881: 44; Wladimiroff 1997). The text is a translation into Yukaghir of the Lord's prayer. Nothing is known about its origin. It is essentially a word-to-word translation and in some respects preserves the Russian syntax, for example, by postposing a possessive pronoun, which is not typical of Yukaghir. The text is written in Dutch orthography; for instance, the digraph *oe* denotes *u*.

Witsen's text was published by Adelung (1806: 562), Schiefner (1871a: 393–394), and Collinder (1940: 14), however these publications omit some diacritic signs and do not provide explanations for some words. Therefore I cite it here as a whole after Witsen (1785: 687). The first line presents Witsen's text, the second line gives possible modern Kolyma Yukaghir correspondences, while the third line provides glosses.

Otje eē: father	Mitsjē, mit-l'ə we-POS	kandi qadi which	koendsjoenga, kužu:-gə sky-LOC	Tcmlælængh ? ?	Nim, nú: name
Totliē, tət-l'ə thou-POS	Legatei l'ə-ŋi-tə-j be-PL-FUT-3	poegandallanpoh, pugu-d-allə-bə sun-GEN-boss-N	Totliē tət-l'ə thou-POS		
Lætiot, l'ə-t-i ət be-FUT-3SUBJ	t'sjemol al čomo:l-əl will-N	kaltei, kel-tə-j come-FUT-3	konda qodo how	koed Zjuga, kužu:-gə sky-LOC	
Je i and	levianh: lebe:-gə earth-LOC	Lünliagel i:-legul smoke-food	Miltje mit-l'ə we-POS	Monidetjelæh, moni-də-tə-lə stomach-TR-FUT-?	
keyck kej-k give-IMP	Mítin mit-in we-DAT	telaman, ?-mə here-TEMP	Je i and	ponkatsj poña:š leave.IMP	Mitin, mit-in we-DAT
Taldelpon tadul-pə debt-PL	Mitlæpoel, mit-l'ə-pul we-POS-PL	Mit mit we	kondan qodo how	poniatsjok poña:š-o:k leave- INTER.1PL	tannevinol tańńə-?-ŋin owe-?-DAT

12 Introduction

Mitlæpoel,	Je	kondo	Olgonílæk	Mitel
mit-l'ə-pul	i	qodo	əl-qanji:-lək	mit-u-l
we-POS-PL	and	how	NEG-chase-PROH	we-0-PRON.ACC

Olo	Oimik,	kondo	molíak	Mitel
əl-?-mik		qodo	moli	mit-u-l
NEG-?-TR.2SG		how	by.no.means	we-0-PRON.ACC

kimda annelan;	Le döt		poegoedal Lenpoh,	Je
kimda:nə-?	l'ə-t	ət	pugu-d-[allə]-bə	i
deceive-?	be-FUT	SUBJ	sun-GEN-boss-N	and

tonbanck,	Je	tændælov,	koendejanck.
tönbə-š-k	i	?	kude-jə-k
strong-V-IMP	and	?	become-?-IMP

As can be seen above, most words from this text have Kolyma Yukaghir equivalents, so we may be dealing with an Old-Yukaghir idiom that was very close, if not identical, to Kolyma Yukaghir. If this is so, the text is likely to have been recorded on the upper Kolyma.

In this dictionary Witsen's materials are indicated as W. The original transcription is preserved, except that, as for the other sources, I write *y* instead of *ï*.

2.1.2. Mueller/Lindenau

Chronologically the next data on Old Yukaghir were recorded by Jakob Lindenau, a participant of the Second Kamchatkan expedition organized by Gerhard Friedrich Mueller.² Lindenau's task was the ethnographic description of the Siberian peoples, including their languages. His Yukaghir materials consist of two wordlists recorded in 1741, when Lindenau traveled from Jakutsk to Oxotsk. Both are translations of the standard Latin list compiled by Mueller, which contains about 300 words from the basic vocabulary. The lists reflect two different Old Yukaghir idioms.

The list I refer to as Mueller/Kolyma (MK) is kept in RGADA in the so-called *Portfeli Millera* (fund 199, list 2, part 513, № 13, pp. 20–23 rev.). Page 20 contains the title: *Jukagirisch (Kolymskije Yukagiri) - Korjakisch - Korjakisch II – Kurilisch*, and consequently some Chukotkan-Kamchatkan words. There are no references to Yukaghirs in the comments provided on page 23, so the exact place of the recording is unknown. It is however clear that we

are dealing with Yukaghir from the upper or middle Kolyma. The list contains about 220 lexical entries, as well as the incomplete verbal paradigm. As far as I know, it has never been published before.

This source is likely to reflect the language of the *Kolym-cy*³ tribe that lived in the middle basin of the Kolyma, as well as on the rivers Jasačnaja and Korkodon (Dolgix 1960). *Kolym-cy* is the Russian word, while the Yukaghirs used the word *kögimə* to refer to this tribe, cf. modern T *kö:jmə* ‘Kolyma Yukaghir’. According to Sauer (1802), Yukaghirs of the upper Kolyma considered themselves the descendants of the people called *konghini*, which suggests the PY **köŋkimə*, cf. also the hydronym *kongina* found in Jochelson’s materials. According to Jochelson (1900: 209), Kolyma Yukaghirs of the end of the 19th century regarded the *kögimə* as an extinct archaic tribe. In the 18th century ten *kögimə* clans merged together on the upper Kolyma into three clans called in Russian Rybnikovskij (Fish clan), Nartennyj (Sledge clan), and Uskansij (Hare clan) (Dolgix 1960: 417). The former later joined the lower Kolyma Yukaghirs and mixed with other tribes (*Soromba* and *Omok-i*), while the latter two became ancestors of the modern Kolyma Yukaghirs. The idiom reflected in MK was apparently spoken by a Northern group of *kögimə*, possibly by members of the Rybnikovskij clan.

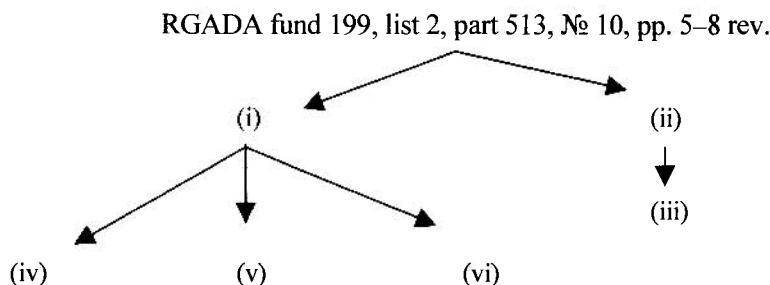
Lindenau’s second wordlist contains about 280 words. It was recorded in Ust’-Janskoe, a settlement on the mouth of the Jana in the middle of the North-Western Yukaghir territory and is abbreviated here as MU (Mueller/Ust-Janskoe). According to Dolgix (1960), in the 17th century this territory was occupied by several Yukaghir tribes: *Jandin-cy* (on the right bank of the Lena), *Xoromo-i* (in the upper basin of the Yana), *Jangin-cy* (around the middle of the Indigirka), *Onojdi* (on the right bank of the Jana), and *Oljubenzi* (in the lower basin of the Indigirka). Veenker (1984) attributed the language of MU to the *Xoromo-i* tribe. However, by the time it was recorded (1741), the North-Western Old Yukaghir tribes had been already greatly mixed. The *Jandin-cy* were a mixed Even-Yukaghir tribe to start with (Dolgix 1960: 382). According to Gurvic (1982), they later mixed with the *Xoromo-i* and *Jangin-cy*, while the *Jangin-cy* mixed with the Evens, in their turn. The tribe *Onojdi* was divided into two clans, Petajskij/Betil’skij and Del’janskij/Zel’janskij. The former was apparently Tungus by origin, and Gurvic (1966: 11), contrary to Dolgix, did not even consider it Yukaghir. So the idiom reflected in MU is likely to be a mixture of several idioms spoken by North-Western tribes, and also had an Even component.

This list is known in several copies. The main copy, arguably written down by Lindenau himself, is preserved in RGADA (fund 199, list 2, part 513, № 10, pp. 5–8 rev.) and contains the following title on page 5: *Lexikon burjatskij, tungusskij, lamutskij* [The Buriat, Even and Evenki vocabularies]. The first

sheet has the additional title: *Yukaghiri Ust-Janskoe* [The Yukaghirs from Ust'-Janskoe]. It is this copy that is published in the present volume and denoted as MU. In addition, there are several copies made by scribes.

- (i) OR GPB, fund 7, № 111, pp. 15–20 rev.
- (ii) AAN, fund 94, list. 1. № 170, pp. 1–14
- (iii) OR GPB, Erm/nem., № 41/1, pp. 95–100
- (iv) OR GPB, Erm., № 577
- (v) OR GPB, fund 7, № 142
- (vi) OR GPB, fund 7, № 111, pp. 21–22

The relationship between these copies can be represented in the following way:



Copy (i) has the following title on page 9: *Tungusike (Werchna Angara) - Tungusice (Jakutz'k) - Yukagiri (Ust-Janskoe)* [Evenki (Upper Angara) – Evenki (Jakutsk) – Yukaghir (Ust'-Janskoe)]. Page 16 rev. contains two additional words absent from the other copies, which I have included in the dictionary: *kakā* 'penis' and *papā* 'vagina'. Copy (i) and consequently all copies derived from it, i.e. (iv), (v) and (vi), demonstrate numerous inconsistencies in rendering the diacritic signs, as well as the following copying mistakes (in order of appearance).

Table 1.

MU	copy (i)	modern Yukaghir	
numă	memă	K numö	house
tschuwónschä	tschiwoonschä	K šubed'ə	heart
jong, jongla	jongjongla	K joyul	nose
anggä	angya	K T aŋa	mouth
tschanngdschamä	tschangdscham	-	wolverine
óllega	olloga	T al'γə	fish

Table 1 cont.

tschul	tschal	K T ču:l	meat, flesh
ljängdschandō	gängdschando	K janʒəd-ō:	(goose) egg
längde-bündsche	lérigder-bundsched	T leŋdəbudʔə	want to eat (INTR.1SG)
längdetsche	längdesche	K T leŋdətɕə	eat (FUT.1SG)
imobujen	imobyjen	K immu:jə	get drunk (INTR.1SG)
ängdsche	anydsche	-	cry
ongetschele	ongetschell	K T oyo:ti:li	stand (FUT.INTR.1PL)
ongénit	ongenet	K T oyo:ŋitəj	stand (FUT.INTR.3PL)
namdátšhit	namdútschit	-	low
lútsche	tútsche	K ju:čə	smoke
-jelaxlon	-jelaklon	K jelekun	four
-andaklon	-andklon	K ataɕun	two
jömbon	jämboon	T jabo:ŋ	dead (INTR.3SG)
ónsche	ousche	K o:zi:	water
vaitagá	baitagá	K pa:jtəgə	woman (AUGM)
nóxtscha	noktscha	K noɕsə	sable
endschdsche	endschöldsche	K endʔə	living (IMPF.PART)
-tschomo	-tschonv	K como:-	big
jehotí	jchotí	K joɣoti:	arrow

Copy (i) or one of its derivatives served as the source of Pallas's publications (1786, 1789),⁴ although Pallas transliterated it into Cyrillic. His materials reflect the deviations from the main copy cited above; for example, he writes *мемя* for *memá*, instead of *numá* (*numo*) 'house'. Notably, Pallas omitted some words from the original list. A few words from copies (ii) or (iii) were published in Schiefner (1859, 1871a) and Veenker (1984).⁵ In this volume the MU list is published in full for the first time.

2.1.3. Klitschka

The materials of the Irkutsk governor Franz (Fedor) Klitschka include the translation of the 22 so-called Bachmeister sentences (*Bachmeistersche Sprachproben*) and numerals. They were sent by Klichka to Pallas in a letter received by the addressee on 10 July 1781. The manuscript is kept in OR GPB (fund 7, № 11, pp. 3 rev. – 5). Page 5 rev. contains the following note: *Réçu le 10 Juillet 1781 de Mr de Klitschka Gouvernier d'Irkutsk*, while page 3 rev. contains the title: *Reči perevedennye srednekovymnskogo jukagirskogo rodu ot knjazca Afanasija Ostrjakova i pročix inorodcev togo že roda* [Sentences translated by the princeling of the Yukaghir Srednekolymsk tribe Afanasij

Ostrjakov and other members of the same tribe]. The so-called Schögren's fund (AAN fund 94, list 2, № 60, p. 5) has a copy of the numerals from Klitschka's materials with the following title: *Perevod na jukagirskoj jazyk, perevodčikov zdes ne slučilos', a nižepisannye slova najdeny v prežnix delax* [A translation into Yukaghir. There are no translators here, but the words below were found in the earlier documents]. So the translation was made in the vicinity of Srednekolym'sk where the Northern group of Kolyma Yukaghirs (*Kolym-cy*) lived. Indeed, from a lexical viewpoint, Klitschka's materials stand very close to Mueller/Kolyma (MK), see 2.1.2.

Klitschka's materials were originally written down in Cyrillic, but published by Schiefner (1871a) in a Latin transliteration. This transliteration deviates from the original at several points, partly because Schiefner attempted to correct supposed errors in the Cyrillic text. The deviations are shown below together with the transliteration used in the present volume and, where possible, approximate equivalents from modern Yukaghir.

Table 2.

Schiefner	KL	my transliteration	K	
el jennili	еленнылыи	el-(j)ennylyi	-	do not want (INTR.1PL)
marxilduek ale amtan	мархилдиекъ алеамданъ	marxildiek ale-amdan	marqil'dö:k əl-amdaj	little girl immortal (INTR.3SG)
cemut	чему	chemu	cumu	everybody
keilani	кейланѣи	kejlani	kejləni	red (INTR.3SG)
manalä	манал'а	manal'a	majlə	hair
oilä	оил'а	oil'a	öj'ə	there is/are no
elejun	элеюч	elejuc	jə:ləj	ill (INTR.3SG)
jojulen	еюлечъ	(j)ejuleč	jə:ləj	ill (INTR.3SG)
pāndai	п'андай	p'andaj	pə:dəj	burn (INTR.3SG)
jonkul	сюнкуль	(j)ejunkul	jəyul	nose
xoni	хонѣи	qoneii	qoni	go (INTR.3SG)
comoc	чомонъ	comon	comə:j	big (INTR.3SG)
juaa	ю'оа	ju'oa	jə:	see (TR.1SG)
lukoc	локонъ	ljukon	juko:j	small
pimzai	нпимжаи-ло	ipimzailo	memzəjələ	flame (ACC)
modoi	модони	modoni	modonji	sit (INTR.3PL)
tin	тыннь	tynn	tiŋ	this
adi	адыи	adyi	adi	firm, hard (INTR.3SG)

Klitschka's materials were published in Jochelson (1900: 229–230), where they were essentially translated into contemporary Yukaghir. In this dictionary they are reproduced in my transliteration after the manuscript and denoted as KL.

2.1.4. Boensing

The materials of the assessor Boensing include the translation of Bachmeister's sentences and some numerals. The original manuscript is kept in OR GPB (fund 7, № 132, pp. 45 rev. – 46 rev.). As follows from the note on page 46 rev., the data were sent to Pallas by Boensing in a letter of 24 March 1781 received by the addressee on 5 September 1781. In addition, the State Public Library in St Petersburg (OR GPB) has several secondary copies in fund 7: (i) № 132, pp. 26–31, (ii) № 132, pp. 32–36, (iii) № 132, pp. 37–41, (iv) № 137, pp. 9–12, (v) № 137, pp. 13–16, and (vi) № 137, pp. 17–20. These contain many copying mistakes and are not considered here. Copies (i), (ii) and (v) as well as the main copy have the following title (the English translation is mine): *Translation of Russian sentences into Koryak, Chukchi and Yukaghir with a separate indication of each language made in the fortress of Giziga according to the printed book sent by Mr. Timofej Šmalev*. Copies (iii) and (iv) have a different title: *Translation from Russian into Koryak, Chukchi and Yukaghir made by the commander of the Okhotsk port Captain Lieutenant Zubov via interpreters*. This indicates that the translation was made in or near Giziga, i.e. not in a Yukaghir territory.⁶

As already suggested by Tailleux (1962: 56–58), the idiom recorded by Boensing is particularly close or even nearly identical to the Čuvan idiom recorded by Matjuškin (on which see 2.1.6 below). This idiom was spoken by the Čuvan-cy tribe, which lived on the Anadyr' and underwent strong assimilation by the Chukchi in the 18th century. The Čuvan-cy were constantly attacked by the Chukchi, especially after 1747 when the Chukchi defeated the local Russian troops and moved closer to Anadyr' (Wrangel 1841: 82; Argentov 1886: 12; Antropova 1957: 179; Okladnikov 1975: 17; Gurvič 1982: 173–74). A part of the Čuvan-cy escaped to Nizhnekolymsk and in the 19th century was assimilated by the Tundra Yukaghirs. Some of them still preserved their ethnic identity at the end of the 19th century, but did not know their language (Nejman 1872: 40; Maydell 1894: 61; Jochelson 1894: 27). Another part of the Čuvan-cy was assimilated by the reindeer breeding Koriaks (Maydell 1925: 24). Finally, the third part settled on the Anjuj in Markovo and underwent russification (see 2.2.5).⁷

Boensing's materials were first published in Schiefner (1871a), while Tailleux (1959a) later reproduced this edition. Apart from omitting many

diacritic signs, the following deviations from the original copy can be seen in Schiefner's publication.

Table 3.

Schiefner	BO	my transliteration	modern Yukaghir	
ili umboti	и'ли' юмботы	ilí-jumboty	T əl-jabatəj	immortal (INTR.3SG)
imoxanbo	имон'хабо	imonxabo	K ńə-yanbo:j	five
tingii	ты'нги	tyngi	T tegi	these
titungec	тн'тһнгѣ'чь	títengēc	-	run (INTR.3SG)
onomikondo	оно'мнкондо	onōmn-kondo	T -köde	man
juxkon	ю'хчонь	júxcoñ	T juko:n	little (INTR.3SG)
makagat	ма'нагатъ	mánagat	K monoyə	how
te	то'	tó	-	eyes
jaizee	я'нже	jánzeje	-	smoke
inza	и'жн'ело	ízhíelo	K o:zi:lə	water (ACC)
elookongat	оно'нгат	onongat	K unungət	from the river (ABL)
uotle-	мотлѣ-	motle-	-	more

In this dictionary Boensing's materials are cited as in the original copy and abbreviated as BO.

2.1.5. *Billings/Merk*

A list of 500 words was recorded during the North-Western expedition directed by Captain Joseph Billings. The place, time and circumstances of the recording are well known from the published diary of the expedition (Sauer 1802). The recording took place on 15 January 1787 in the settlement *Nunkuntung-nagel* (K *nungəd-ən-añil'*) situated in the mouth of the river Nelemnaja, a tributary of the Jasačnaja. These materials also reflect the language of the *Kolym-cy* tribe (see 2.1.2). All words found in Billings' materials are known in modern Kolyma Yukaghir or the earlier variety of Kolyma Yukaghir recorded by Jochelson.

Several copies of the list have survived, but they show significant differences.⁸ Three copies are based on the English orthography: (i) Billings' copy (OR GPB, fund 7, № 137, pp. 23–30 rev.); page 30 rev. contains a note: *Aus Billings Papieren*, (ii) the publication of Sauer (1802), and (iii) the

publication of Schiefner (1871a). Further, there are two publications in Cyrillic, Saryčev (1811: 93–102) and Pallas (1789). They both represent the transliteration of the English-based list, but lack several entries. The transliteration of the digraphs *tsh*, *zsh*, *sh*, *dzsh* as well as *gh* by Saryčev is inconsistent and there are other mistakes that may have appeared when transliterating from Latin to Cyrillic script. These are shown below.

Table 4.

Sarycev	Sauer	K	
анщү	antoo	ad-ö:	son
ларкунъ	larkul	larqul	root
ниченне	natshennee	na:cəñi	sharp (INTR.3SG)
онманнсы	onmannee	önməni	clever (INTR.3SG)
милглiалхлонъ	malgialachlon	malʏi-jalo:j	eight (INTR.3SG)
аньнiома	angnuma	aŋnumə	earlier
неионмонгуль	nejunmolgul	nə-molyil	year
хунишки	chuniirki-	kunir-kil'd'o:j	nine (INTR.3SG)
еленджiенъ	ellendzshien		
умат	umat	jö:mət	see (TR.2PL)
iакъ	i-ak	aja:k	glad (IMP)
этхя	etchea	ece	father
ыаду	yādoo	-	husband
иви	ivi	ajbi:	shadow
нондри	nondri	nonzo:j	weak (INTR.3SG)
моiпохаia	moinchaija	-	wave
юкенъ	irken	irkin	one

Both Saryčev's copy and Sauer's list contain common copying mistakes absent from other sources, cf.:

Table 5.

Sarycev	Sauer	K	
лукундай	lukundae	jukud-ö:	boy
маналла'	manalla'e	majlə	hair
ичеэндамен	itsheendamey	i:cəd-amun	elbow
пундалволе	poondalvolle	po:d-olbəl'ə	widower
лиеуша	lie'usha	ju:čə	soul, spirit
аррангiя	arra'ngia	arannəj	light (INTR.3SG)
iоангуичель	i-oanguitshel	jo:n-kicil	forehead

Table 5 cont.

Billings	Schiefner	K	
lukundä	-	jukud-ð:	boy
-	monolä	majlə	hair
itsheendami	itscheendami	i:cəd-amun	elbow
poondolvole	poondolvole	po:d-olbəl'ə	widower
lewsha	lewsha	ju:cə	soul, spirit
arrangya	arannei	aranñəj	light (INTR.3SG)
i-onguitshel	ionguitshel	jo:n-kicil	forehead

But the following words show that Saryčev did not copy directly from Sauer, since Sauer's list contains copying mistakes not followed by Saryčev.

Table 6.

Saryčev	Sauer	K	
мархель	marchet	marqil'	girl
лагетак	sagetak	legitək	feed (IMP)
пуданмаи	pudanniai	pudenme:j	tall (3SG)
онманнсы	onmanneig	onmāni	clever
поинюлечь	poinjuletsk	pən-jouləc	evening
маагиль	māājil	mayil	coat
кундшу	kunshu	kuzhu:	sky

This indicates that Saryčev and Sauer used a common source based on English orthography. Pallas's publication contains the same Cyrillic transliterations as Saryčev's list, but also shows the individual mistakes illustrated below.

Table 7.

Pallas	Saryčev	Billings	K	
суута	-	tshukta	suktə	city; Jakutsk
аачель	-	aktshel	ekcil'	boat
агангня	-	alangnae	elenñə-	lazy
ягмак	ягтак	-	jaqtək	sing (IMP)
аныюлма	аньниома	angnuma	aŋnumə	earlier

Some of these mistakes could have emerged as a result of the misunderstanding of the Cyrillic form, which suggests that Saryčev and Pallas used a Cyrillic-based source.

On the other hand, Billings' and Schiefner's materials exhibit common features and contain considerably more words than those of Saryčev and Sauer, although they lack a number of words present in the latter. The following table demonstrates copying mistakes shared by Billings and Schiefner.

Table 8.

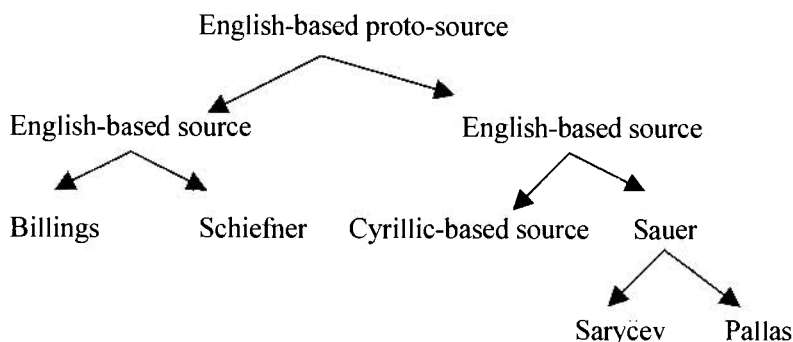
Saryčev	Sauer	Billings	Schiefner	K	
те	tia	-	-	ti:	here
канбунной	kanbunnoi	kanbonnoi	kanbonnoi	kenbunəj	wide (INTR.3SG)
ханнинь	cha'nnin	channen	cha'nnen	qanin	when
-	nunbur	nunbar	nunbar	-	dew
кейвей	keivey	keivy	keivy	kejbəj	thin (INTR.3SG)
онджиль	ondzshil	onzshil	onzhil	ud'il'	nail
онджн	ondzshi	onzshi	onzshi	o:zi:	water

Billings' list also contains some individual features not shared by Schiefner.

Table 9.

Billings	Schiefner	K	
marklōo	marchloo	marql'-ō:	daughter
oniak	aniak	aññək	speak (IMP)
nengashe	nenganshe	nigizə	yesterday
pomzshole'ni	pomzsholene	pōmzələñi	round (INTR.3SG)
angunma	angnuma	aŋnumə	earlier
tshayoia	-	coyojə	knife

This suggests that Billings' and Schiefner's lists both go back to another English-based source, presumably different from the one used by Saryčev and Sauer (see Veenker 1984: 573 for a similar conclusion about Schiefner's and Pallas's publications). In other words, the inheritance schema for this group of sources can be reconstructed in the following way.



In this dictionary I cite the reconstructed forms of the English-based proto-source and indicate them as B. The reconstructed forms are cited with an asterisk if they do not coincide with any of the actual sources. Their relationship to the individual sources is shown in Table 10.

Table 10.

reconstruction		Billings	Sauer
*aimaivi	black	aimaivi	aimáibi
*artshetshúngzsha	sorrow	arlshetshunzsha	artshetshünzsha
*arráni	light	arrangya	arrángiä
*etcheä	father	etschea	etchea
*illeyennee	wind	illeyennee	illejennie
*i-onguitshel	forehead	i-onguitshel	i-oanguitshel
*kailey	raven	bailey	-
*ke-ick	give	ke-ich	keick
*koikil	stump	-	koikél
*ledemnee	low	ledemnee	ledemnie
*poldshitsha	leaf	-	paldshitsha
*poóndolvölle	widow	poondolle	poondalvolle
*yakandaitsa	horse	-	-
*yālon	three	yalon	iālon
*yārka	ice	yarka	iārka
*yelaklon	four	yelaklon	ielahlon
*yelondzshendigia	sun's ray	yelonzshendigia	jelondshent digia
*yelongedukshimba	east	yelongidukshimba	jelongedukshimba
*yomüel	neck	yomuel	jomüel
*yòatsh	pain	yoatsh	joatsh
*yonrul	mow	yourul	-

Table 10 cont.

reconstruction		Saryčev	Schiefner
*aimaivi	black	аймайви	aimaivi
*artshetshúngzsha	sorrow	арчечунжа	artscha tschungzsha
*arranei	light	аррангя	arannei
*etcheā	father	этхя	etcheā
*illeyennee	wind	иллеънье	illevennee
*i-onguitshel	forehead	іоангуйчель	ionguitel
*kailey	raven	-	-
*ke-ick	give	кеик	-
*koikil	stump	коикиль	-
*ledemneē	low	ледемне	ledemnee
*poldshitsha	leaf	полдишча	peldshitsha
*poóndolvolle	widow	пундалволе	poondolvolle
*yakandaitsha	horse	-	yakandaitscha
*yālon	three	ялонъ	yalon
*yārka	ice	іарка, ярка	yarka
*yēlaklon	four	ялаглонъ	yelaklon
*yelondzshendigia	sun's ray	-	yelonzshendigia
*yelongedukshimba	east	іелонгі	jolongodukshimba
*yòatsh	pain	іоачъ	joatsh
*yomūel	neck	іомуель	jomuel
*yonrul	mow	-	jonrul

Finally, three wordlists that resulted from the work of the same expedition are based on German orthography: (i) OR GPB, fund 7, № 137, pp. 2 rev. – 8; page 2 rev. contains the title: *Bey de Billingschen Expedition ven Dr. Merk gesammelt*,⁹ (ii) AAN, fund 94, list 2, № 59, pp. 1–8; page 8 contains the following note: *Captain Billings and some explanation about the Yukaghir calendar in English*, as well as the title: *Vocabulary of the Dialect of the Kovima Ukagers*, and (iii) the publication of Rohbeck (Billings 1802: 91–129). Copies (i) and (ii) are almost fully identical, except that in (ii) all words begin with a capital letter, as was usually the case when copies were made by a scribe. Copy (iii) is a transliteration from English into German orthography. Its source was either Saryčev or a Cyrillic-based list that served as Saryčev's source. This can be seen from the fact that the word *онманнсы* 'wise' in Saryčev's list appeared as a result of the miscopying of the word *onmannei*, while Rohbeck re-transliterated it as *onmannsy*. The German-based list (i) is cited in the dictionary as ME. Since none of the mentioned sources contains all the words found in the other source, the

present volume is the first publication of Billings and Merk's materials in full.

Whether B and ME had a common source remains unclear. The words in these sources are written in the same order, but they are likely to result from parallel recordings of the same speaker by two different members of the expedition. This could account for the divergences in citing grammatical forms. For example, the entry 'to see' is represented as the 2nd person Plural in B (*umat*) and as the Imperative and the 2nd person Plural in ME (*juk* and *jumat*); the entry 'thou' is translated as the Nominative in B (*tat*) and the Predicative in ME (*tattak*), and the entry 'to eat' is represented as the Action Nominal in B (*landgal*) and the Imperative in ME (*langdak*, *lagk*). Such divergences could have appeared if the informant repeated the word several times in different grammatical forms and each recorder wrote down a different form.¹⁰

2.1.6. *Matjuskin (Čuvan and Omok)*

Further Old Yukaghir materials were recorded by sub-lieutenant Fedor Matjuskin, a participant of Wrangel's expedition (1821–1824). The recording was made in 1821 on the river Malyj Anjuj (Wrangel 1841: 81). Matjuskin's materials represent two Cyrillic-based word lists, the so-called *Čuvan* list (about 210 words) and the *Omok* list (134 words), plus five *Čuvan* sentences. Compared to other Old Yukaghir materials, the quality of the transcription is very poor and both lists, especially the *Omok*, contain numerous mistakes. A large number of words cannot be identified, probably due to fundamental mishearings or misspellings. The original manuscript seems to be absent from the materials of Wrangel's expedition kept in the State Navy Archive in St Petersburg. The lists we have at our disposal were first published in the diary of the expedition (Wrangel 1841: 115–125) and reproduced in Starčevskij (1889: 421–437, 461–464, 639). The latter version contains many copying errors, but served as the basis of some phonological conclusions in Angere (1956), see Tailleux (1962: 57).

The *Čuvan* list reflects the language of the Markovo *Čuvan*-cy (see 2.1.4). The identification of the *Omok* materials is difficult because the term *omo* is present in all varieties of Yukaghir.¹¹ Both in modern Kolyma and Tundra Yukaghir it generally means 'people, tribe', cf. also K *om-ni*: 'people'. It remains unclear to which of the Yukaghir tribes Matjuskin referred as *Omok*. What is certain is that the *Omok* materials exhibit phonetic parallels with Tundra Yukaghir not shared by other varieties of Yukaghir (see 4.2.3 and 4.2.4). However, there are no reasons to believe that the *Omoks* were direct ancestors of all the modern Tundra Yukaghirs; rather they may constitute

one of their components (Nikolaeva, *forthc.b*). In fact, Tundra Yukaghir exhibits more Even loanwords than the language recorded by Matjuškin and differs from it in several other respects.

Tailleur (1959a, 1962) was the first to analyze Matjuškin's materials. We owe him the identification of many *Cuvan* and *Omok* words, although in many cases his solutions differ from mine. In the present volume I cite the *Cuvan* and *Omok* materials in transliteration after Wrangel (1841), and indicate them as MC and MO, respectively.

2.2. End of the 19th century and 20th century

2.2.1. Kurilov

The modern Tundra Yukaghir data denoted here as T are cited after the works of Gavril Kurilov, mostly his *Yukaghir-Russian dictionary* (Kurilov 2001) and its shorter version (Kurilov 1990). Lexical entries included in the dictionaries are taken from Kurilov's large collection of texts, some of which have been published in Latin-based transliteration in Maslova (2001) and in Cyrillic in Kurilov (2005). These data are transliterated as described in 1.4. I have tried to cover these sources as fully as possible, but did not include all productive derivational forms cited in Kurilov's dictionaries. Sometimes my English translations differ from his Russian ones.

Tundra Yukaghir has (or until recently had) two local idioms, *Qanaji:* and *Alaji:*, which exhibit slight phonetic differences (Kurilov 1987: 5–8; Krejnovič 1958: 24–25). Kurilov mostly recorded the *Alaji:* variety originally typical of the so-called First Alazeja tribe on the lower Alazeja. More precisely, his materials reflect Tundra Yukaghir as spoken in the last third of the 20th century in the settlement Andrjuškino. *Qanaji:* Yukaghir was spoken by the clans who nomadized between the lower Kolyma and Alazeja and is primarily known from the works of Jochelson (2.2.4).

2.2.2. Krejnovič

Eruxim/Evgenij Krejnovič's materials are indicated in this dictionary as KK for Kolyma Yukaghir (Krejnovič 1982) and KT for Tundra Yukaghir (Krejnovič 1958, 1982). The Cyrillic transcription is transliterated on a regular basis. However I do not indicate some non-phonemic

pronunciational variants such as *ɹ*, which, according to Krejnovič (1982: 12), represents the stop variant of the uvular fricative *ʕ* (i.e. *ɣ*).

Krejnovič made his recordings approximately between 1940 and 1980. They mainly reflect the *Alaji*: variety of Tundra Yukaghir.

2.2.3. *Spiridonov*

The materials of Nikolaj Spiridonov are denoted as SD and reproduced after Spiridonov (2003). This work is a publication of Spiridonov's manuscript Yukaghir-Russian and Even-Russian dictionaries edited by Aleksej Burykin. The original manuscripts have not been preserved. The dictionaries are known from the copies made by the ethnographer E. Orlova, which are currently kept in the Magadan Regional museum.

Spiridonov himself was a native speaker of Kolyma Yukaghir. The Yukaghir part of the dictionary reflects the language of the Kolyma Yukaghirs of 1930–1931. According to the editor of the published version, the dictionary remained unfinished: there are considerably more words beginning with the letter *a* than those beginning with the last letters in the alphabet. The dictionary includes about 750 entries, and many of them are toponyms and hydronyms. The published version contains many misspellings, as the text was copied several times by people who did not have a good knowledge of the language, first of all Orlova herself.

Spiridonov used a mixed orthography based both on the Cyrillic and Latin writing systems, as had been customary in some publications on Northern languages in the early 1930s before the unified writing system was adopted. The 2003 publication omits the diacritics that apparently indicated stress. In this volume the Cyrillic letters are transliterated.

2.2.4. *Jochelson*

Woldemar Jochelson's materials reflect the varieties of Yukaghir spoken at the end of the 19th century and the very beginning of the 20th century. He recorded Kolyma Yukaghir and the *Qaqaji*: variety of Tundra Yukaghir.

Jochelson's materials are represented by two types of data. First, I have used his published works (Jochelson 1898, 1900). They are marked as KJ and TJ for Kolyma and Tundra Yukaghir, respectively. Jochelson's other published works (e.g. Jochelson 1905) are not used, as they hardly contain any words absent from the publications mentioned above. In rendering these Cyrillic data I have followed my usual transliteration practice with the following amendments:

Jochelson's \bar{a} is rendered as e , d_e , which optionally represents \bar{z} , is rendered as \bar{z} , μ is rendered as d' , and the sign indicating the optional palatalization on \bar{c} is omitted. Stress is not shown.

Second, I have used Jochelson's manuscript dictionaries, which remain largely unpublished. They are kept in the form of card files in the Archive of St Petersburg's Institute of Oriental Studies in the so-called *Jochelson's Fund*.¹² The Kolyma Yukaghir dictionary is found in fund 23, part 1, № 17–22 under the title *Yukaghir-Russian dictionary*. Part of this dictionary was published in Jochelson (1926: 317–327), but the manuscript is much more substantial. In the present volume these materials are denoted as KD. The Tundra Yukaghir dictionary is found in fund 23, part 1, № 13–16 under the title *Khangai-Russian dictionary*. A few *Qanqaji* words are cited in Jochelson (1926: 317–327). In the present volume these materials are marked as TD. I also cite Yukaghir words from the text of Jochelson (1926) and have included them in KD and TD.

Although my publication of the manuscript dictionaries is more complete than Jochelson (1926), I have not reproduced them in full. A separate edition of these would be an interesting project, though not strictly required for the purpose of this dictionary. In particular, I have omitted many idiomatic expressions and sentence examples, as well as Jochelson's comments on the origin and usage of some words. In the manuscripts verbs are cited in the Imperative and in the 3rd person Singular forms, while I only cite verbal stems. In the Khangai dictionary Jochelson cites nouns in the emphatic form in $-ŋ$, but I cite them uninflected. The dictionaries are written down in Latin-based transcription. While rendering it I made the following additional changes: stress is not indicated, l stands for l and l' stands for l , \acute{n} stands for n , h for g . (both h and g represent non-phonemic variants of γ in the original), η for \bar{n} , d' or \bar{j} , \bar{z} for d_e , and m for m' . In the original, m' indicates a non-phonemic palatalization before the diphthong e : [ie].

2.2.5. Schiefner

Anton Schiefner was the first to undertake a linguistic study of Yukaghir, in the second half of the 19th century. In his three articles (Schiefner 1859, 1871a, 1871b) he provided short grammatical notes and published word lists compiled from previous archival sources, as well as the new data he himself collected through correspondence with the local administration. On Schiefner's publications of earlier sources see 2.1.2, 2.1.3, 2.1.4 and 2.1.5.

Schiefner's materials come from the following sources. First, in 1859 he received two short Yukaghir texts from the governor of Jakutsk, Dr. Julius

Stubendorff, and one text from the teacher in a Jakutsk primary school, Fedor Rajsikj. These texts were recorded from an Omolon Yukaghir who visited Jakutsk in 1858 and represent the same idiom. Schiefner published them twice, in Schiefner (1859) and, with the improved transcription and translation, in Schiefner (1871a). In this dictionary I cite these materials after the latter publication and denote them as RS. Second, in 1861 Schiefner received a list of 2289 words, 50 sentences and one short text from Petr Suvorov, the Chukchi missionary in Srednekolymsk. Suvorov's materials are published in Schiefner (1871a). In this dictionary they are denoted as SU. RS and SU reflect the idioms spoken on the Omolon and in Srednekolymsk, respectively. They represent geographically and linguistically close varieties of Kolyma Yukaghir.

Third, Schiefner (1871b) presents 111 sentences recorded by the General Governor of Eastern Siberia, Baron Gerhard von Maydell. These are indicated here as M. Maydell himself was involved in geographic and ethnographic work in North-Eastern Siberia in 1868–1870 (see Maydell 1894–1896, 1925). He recorded his Yukaghir materials in 1870 on the river Anadyr' in the settlement Markovo. The basin of Anadyr' and Markovo in particular was then inhabited by the Yukaghir *Čuvan*-cy tribe, see 2.1.4. However, Maydell noted that most *Čuvan*-cy could not speak their language when he visited them. Some had begun to speak Russian, while others had been assimilated by the Kolyma Yukaghirs (Maydell 1894–1896; Jochelson 1905; Gurvič 1966: 142). Maydell found it difficult to find speakers of the language in Markovo. Only one elderly lady was able to provide him with translations of Russian sentences into Yukaghir, and the influence of Russian syntax is evident in these translations. These materials do not show significant linguistic differences from Kolyma Yukaghir and hardly represent the language of the *Čuvan*-cy.

Most of Schiefner's materials were re-published in Jochelson (1900: 227–229, 231–237), however Jochelson made many changes and corrections, essentially adapting them to the language spoken at the end of the 19th century. I cite Schiefner's materials after his publications, with the following changes in transcription: *ŋ* instead of *ñ*, *d'* instead of *ʒ* ('), and *γ* instead of *ǵ*.

3. Basics of Kolyma Yukaghir phonology

In the dictionary I have used my own phonological transcription for modern Kolyma Yukaghir. Here I explain this transcription and provide a phonological description of the language, which serves as the basis for the Proto-Yukaghir reconstruction in Section 4. This description should not be viewed as comprehensive: it only concentrates on certain relevant topics. Note that it deviates from previous descriptions (Krejnović 1982; Maslova 2003) on a number of points. In some cases I have chosen to compare the Kolyma Yukaghir data with modern Tundra Yukaghir. In this Introduction I have rendered the Tundra Yukaghir in my phonological transcription, although in the text of the dictionary I have left all data other than the Kolyma Yukaghir as they appear in the existing sources. Where in this Introduction there is no indication of the source, the examples are taken from Kolyma Yukaghir.

3.1. Phonemic inventory

3.1.1. *Vocalism*

Kolyma and Tundra Yukaghir have identical vocalic systems as below.

(1)	front		back	
	unrounded	rounded	unrounded	rounded
high	i i:		u u:	
non-high	e e:	ø ¹ ø:	a a:	o o:

Short non-high vowels are not specified for place of articulation when they occur after the first bimoraic foot, i.e. after the initial structures CVCV, CVC or CV:. As was first suggested in Nikolaeva (1986), in this position they are represented as a neutral vowel ə, which may partly harmonize to a full vowel (see 3.3.2). The vowel ə is not included in the phonemic inventory, because its distribution is fully predictable from the syllabic structure. Therefore it does not contrast with full vowels. However, I find it useful to indicate it in the transcription because, as shown below in 3.4.1 and 3.4.2, ə has distinct phonological properties not shared by other vowels. The vowel ə is also present in a few monosyllabic functional words that normally form a single phonological phrase with the adjacent lexical word,

such as the pronouns *mət* ‘I’, *tət* ‘thou’, and possibly some forms of the auxiliary verb *lā-* ‘to be’.

For Tundra Yukaghir Kurilov usually writes diphthongs instead of long mid vowels, but not always consistently, cf. T *xoolew-* ‘to kill’ (Kurilov 1990: 307) and *xuole:w-* id. (Kurilov 1990: 308). There are variations in Jochelson’s and Krejnović’s materials, too, e.g. KJ *xamluo-* ~ *xamlo-* ‘how many, how much’ and KK *juo-* ~ *jo:-* ‘to see’. Krejnović (1958: 9, 1982: 9–10) explicitly mentioned the diphthongs *ie* and *uo*, but did not discuss their phonological status. However, there are no minimal or quasi-minimal pairs that can distinguish diphthongs from the corresponding long mid vowels. Instead, the falling diphthongs are the non-phonological variants of long mid vowels, i.e. *e:* can be realized as *ie*, *o:* as *uo*, and *ö* as *uö* (or *uo*).² Diphthongs are generally preferred in the stressed position, especially in monosyllabic words, but this distribution is a tendency rather than a strict rule. There are only two words in Kolyma Yukaghir where the long *e:* never diphthongizes: *me:me:* ‘bear’ and *emme:* ‘mummy’. The former is apparently an emphatic taboo word which appeared as a result of reduplication (see 4.3), while the latter is an affectionate nursery word. T *ne:nuke:* ‘riddle’ and *eke:* ‘elder sister’ do not seem to allow diphthongs either, but both are recent Even borrowings. In any case, these few words do not provide sufficient evidence for the phonological status of falling diphthongs.

Maslova (2003) analyzes the glides *j* and *w* in the syllable-final position as consonants. In this Introduction a different solution is adopted: the glides *j* and *w* before a consonant or a pause are analyzed as vocalic components of rising diphthongs and are transcribed as *j* and *u*, respectively. There are no prevocalic rising diphthongs, but the glides are realized as the consonantal *j* or K *b* ~ T *w*, cf. *moj-m* ‘holds (TR.3SG)’ vs. *moj-o:-j* ‘is held (RES-3SG)’, *kou-dā-* (optionally [*kōwdā-*]) ‘to drive out’ ~ *keb-ej-* ‘to leave’, and T *lau-m* ‘eats (TR.3SG)’ (optionally [*lawm*]) ~ *law-i-tā-* ‘to feed’.

The following arguments support this solution. First, it follows from the phonotactic properties of the language. According to Maslova (2003: 34), consonantal clusters are disallowed in the syllable-final position with the exception of *j*-initial clusters, e.g. *qojl* ‘God’ and *qojl-gət* (ABL). The diphthong analysis strengthens the generalization: if *j* is analyzed here as vocalic, such words do not present exceptions. Similarly, in Kolyma Yukaghir the Imperative morpheme *-k* attaches after a vowel-final stem, but is absent after a consonant, since a word cannot end in two consonants (see 3.6), cf. *mada:-k* ‘sit down (IMP)’ but *sörilās* ‘write (IMP)’. But *-k* occurs after *j*, e.g. *kej-k* ‘give (IMP)’, apparently because the latter is vocalic. Second, the glide can be followed by the epenthetic *l*, which only occurs

between two vowels (3.6.1), cf. *tadi:-l-a:-* ‘to start giving’ (INCH) and *uj-l-a:-* ‘to start working’ (INCH). Third, as discussed in 3.4.1, in monosyllabic nouns the vowel is always long. This is because such words must be bimoraic, but the final consonant does not provide a mora. However, monosyllabic nouns ending in *u* or *j* do not contain long vowels, e.g. *jou* ‘illness’ and *köj* ‘man’. This indicates that the final segment here is not a consonant but a moraic vowel. Finally, the consonant *j* does not occur at the end of a consonantal cluster (see 3.2). As noticed in Maslova (2003: 34), the only exception is the cluster *wj*, e.g. *šəwjemet* ‘came in (INTR.2PL)’ (in Maslova’s transcription). With the analysis of *w* as vocalic, this case does not present an exception: *j* here functions as a syllable onset. I transcribe this word as *šəujəmət*.

The rising diphthongs in the first foot can contain any vowel, e.g. *köj* ‘man’, *pajpə* ‘women’, *pujm* ‘blows (TR.3SG)’, *tolou* ‘wild reindeer’, and *leu* ‘eat (TR.1SG)’. After the first foot only two diphthongs involving non-high vowels seem to be allowed in non-compound words, *ej* and *aj*. Their distribution depends on the harmonic quality of the stem in the same way as the distribution of harmonizing long vowels *a:* and *e:* (see 3.3): *ej* occurs in front stems and *aj* occurs in back stems. For example, these diphthongs are present in certain verbal affixes such as the Perfective *-ej/-aj-*, *-dej/-daj-*, *-rej/-raj-*, *-tej/-taj-*, and *-šej/-šaj-*, e.g. *ločil-daj-* ‘to make fire’ vs. *šel’gə-dej-* ‘to break’, *čoly-aj-* ‘to push’ vs. *jed-ej-* ‘to appear’. Note that although the only short non-high vowel allowed after the first foot is *ə*, these affixes do not contain *ə*. As is shown in 3.3.2, *ə* optionally harmonizes in rounding to the vowels of the first foot, so if they contained *ə* we would expect the form **ločil-doj-*, at least in some idiolects. However, such forms do not exist. This provides an additional argument for analyzing *ej* and *aj* as diphthongs, rather than the combination “short vowel + *j*”.

3.1.2. Consonantism

The consonantal systems of Kolyma and Tundra Yukaghir differ slightly. The table below marks with the symbols T or K the phonemes that are particular to either language.

(2)	labial	coronal	palatal	velar	uvular
voiceless stops	p	t		k	q
voiced stops	b	d		g	
voiceless fricatives		T s, K š			ɣ
voiced fricatives		K ž			
voiceless affricates			č		
voiced affricates			dʲ		
nasals	m	n	ɲ	ŋ	
laterals		l	ɭ		
trills		r			
approximants	T w		j		

Krejnović (1982) records the palatal stop *tʲ* instead of *č*, however in modern Kolyma Yukaghir it is pronounced as a palatal voiceless affricate and is so transcribed by most other sources.

According to Maslova (2003), *b* is not a phoneme in Kolyma Yukaghir, which instead has the phonemic *w*. She takes *b* to be a variant of *w*, either a free variant in the intervocalic position or a positional variant after a sonorant. However, the alternations shown by *b* are similar to the alternations of other voiced stops. It is realized as *b* after a sonorant and as *b* or *w* (in some idiolects) intervocalically. In the coda position *b* undergoes assimilative processes typical of other voiced consonants (see 3.5.2). For this reason I have treated *b* as a voiced stop in this volume. So, unlike in Maslova's description, all voiceless stops have a voiced counterpart: *p* ~ *b*, *t* ~ *d*, *k* ~ *g*, *č* ~ *dʲ*, *š* ~ *ž*, and *q* ~ *ɣ* (in the latter case there is also a difference in the mode of articulation).

As for *w*, I have argued in the previous subsection that in Kolyma Yukaghir it does not have a phonemic status. Rather it is an optional pronunciational variant of *u*, which occurs before a consonant as the second component of a diphthong. Before a vowel, *w* may optionally replace *b* in some idiolects. In contrast, in Tundra Yukaghir *w* is phonemic: it occurs word-initially and contrasts with *b* in the intervocalic and syllable-initial position, e.g. T *qabaŋa*: 'bald' vs. *qawarqə* 'pit', *ilwi*:- 'to graze' vs. *tilba*:- 'to trample down'. The phonemes *w* and *b* in Tundra Yukaghir have a different origin, see 4.2.

The distribution of *k/q*, on the one hand, and *g/ɣ*, on the other, is not free but depends on the harmonic quality of the stem (see 3.3.4). But since the domain of synharmonism does not include inflectional suffixes, one can easily find minimal and quasiminimal pairs distinguished by velar vs. uvular

consonants, e.g. *monoya* ‘mountain sheep’ vs. *nono-gə* ‘thorn (LOC)’. This suggests that *k*, *q*, *g*, and *ɣ* are distinct phonemes.

3.1.3. Some non-phonemic variations

Most allophonic variations are described in Krejnovič (1982) and Maslova (2003); here I will only discuss the variations relevant for my transcription. Voiced stops tend to undergo fricativization in the intervocalic position. For example, Krejnovič and Maslova normally write *ś* as the allophonic variant of *č*. *ś* occurs intervocalically or word-finally (3a) in contrast to the *č* available word-initially or after some consonants (3b). But I write *č* in both cases, on phonological grounds.

- | | | | |
|-------|--------------|---------|------------|
| (3)a. | <i>ece:</i> | [eśie] | father |
| | <i>kebeč</i> | [kebeś] | went (3SG) |
| b. | <i>erčə</i> | [erčə] | bad |
| | <i>čobul</i> | [čobul] | sea |

Another example is the inter-speaker variation *b* ~ *w* mentioned in the previous subsection, which is reflected in Maslova’s transcription. Such variations are fully predictable and are not recorded in my dictionary.

The previous sources have not mentioned that the vowels *a* and *o* undergo optional fronting after the palatal consonants *j*, *č*, *lʲ*, and *ň*, especially if they are unstressed, e.g. *janmägej-* ~ *jenmägej-* ‘to run into’ and *ňajdo:-* ~ *nejdo:-* ‘spare’. These variations are reflected in my data. The initial prevocalic *j* can be deleted, which leads to various changes, e.g. *jomil* ~ *imil* ‘neck’, *ju:-legul* ~ *i:-legul* ‘flour’, *jurgud’e:jə* ~ *irgud’e:jə* ‘star’, *jöulugə* ~ *i:lugə* ‘sorrow’, and *joŋžo:-* ~ *iŋžo:-* ~ *uŋžo:-* ‘to sleep’. In other instances the initial *j*, *lʲ* or *ň* assimilates to the second syllable, e.g. *je’lo:d’ə* ~ *le’lo:d’ə* ‘sun’, *ňumušej-* ~ *mumušej-* ‘to press’, *la:jə* ~ *ja:jə* ‘spleen’, *jouluč-* ~ *louluč-* ‘to ask’, and *ja.zi:-* ~ *ňa.zi:-* (< **ňančə-*) ‘to scratch’.

Other non-predictable optional variations that may be reflected in my transcription are *rqlʲ* ~ *ślʲ* (e.g. *maslo:* < *marqlō:* ‘daughter’), *ql* ~ *qś* (e.g. *ataqləštə* ~ *ataqśəštə* ‘second’), *m* ~ *b* (e.g. *mi:bə* ~ *mi:mə* ‘basis, character, custom’), *ö* ~ *u* (e.g. *lukil* ~ *lökil* ‘blunt arrow’), *i* ~ *u* (e.g. *mimil* ~ *mumul* ‘young antlers’), and *i* ~ *e* (e.g. *mi:d’i:* ~ *me:d’i:* ‘sledge’).

3.2. Positional restrictions on consonants

Since the syllabification constraints do not allow more than one consonant in the coda or onset position (see 3.6), Kolyma Yukaghir only exhibits two-consonantal clusters. A table of attested clusters is presented in Krejnovic (1982: 31), but it appears to miss some generalizations, namely, that consonants fall into several natural classes based on their combinatorial properties: voiceless obstruents (stops, affricates, and fricatives), voiced obstruents, the vibrant *r*, other sonorants, and *j*. It also fails to mention that the morpheme boundary plays a certain role in the well-formedness of clusters either. Table (1) represents possible combinations across the morpheme boundary. Table (2) shows the co-occurrence of consonants within a morpheme. The symbol # stands for the word boundary.

Table 1.

	voiceless obstruents	voiced obstruents	sonorants	<i>j</i>	#
voiceless obstruents	+	-	+	-	+
voiced obstruents	-	-	-	-	-
sonorants	+	+	+	-	+
<i>j</i>	+	+	+	-	+
#	+	-	+	+	

Table 2.

	voiceless obstruents	voiced obstruents	<i>r</i>	other sonorants	<i>j</i>
voiceless obstruents	+	-	-	-	-
voiced obstruents	-	-	-	-	-
<i>r</i>	+	+	-	+	-
other sonorants	-	+	+	+	-
<i>j</i>	+	+	+	+	-

As follows from the tables, clusters ending in *j* are totally forbidden. The syllable-initial *j* assimilates to the previous consonant both across the morpheme boundary and inside a morpheme (3.5.4). Voiced obstruents only occur intervocally, after a sonorant or a glide, and alternate with other consonants elsewhere (3.5.1 and 3.5.2). They are disallowed word-initially with the exception of a few words, but in Tundra Yukaghir word-initial voiceless consonants may undergo voicing (3.5.3).

Generally speaking, the combinatorial potential of a consonant within a morpheme is lower than across a morpheme boundary. Sonorants do not normally precede voiceless obstruents within a morpheme due to the historical process of voicing (see 4.2.5). However, there are no restrictions at the morpheme boundary, e.g. *ludul-tə-* ‘to provide with iron (iron-TR)’ and *urun-tə-* ‘to provide with a bed (bed-TR)’. The consonant *r* typically co-occurs with tautomorphemic voiceless obstruents, but in some cases it may precede a voiced obstruent.

There are additional restrictions on the individual consonants not reflected in the tables. The consonant *ŋ* does not occur word-initially and *ŋ* is not allowed word-finally, although it is possible in Tundra Yukaghir. In Kolyma Yukaghir *ŋ* is depalatalized if it appears at the end of a word, e.g. *kejbən* ‘make thinner (IMP)’ from *kejbəŋ-* ‘to make thinner’. The clusters *lŋ* and *jl* usually assimilate into *ŋŋ* and *jj*, respectively. This creates geminates, which are otherwise rare. The clusters *nl* and *nl'* are generally disallowed, at least at the morpheme boundary. In such clusters *n* assimilates to *l* with further palatalization before the palatal *l'*. Examples are *kil-lə* ‘whose’ < *kin* ‘who’ + *lə* (Possessive) and *mol-ləl* ‘said (EV.INTR.3SG)’ < *mon-* ‘to say’ + *ləl* (Evidential). But when a sonorant-final stem is followed by an *l*-initial affix, *l* usually falls out, e.g. *nugenə* < *nugen-lə* ‘hand (ACC)’ and *qa:rək* < *qa:r-lək* ‘skin (PRED)’. The geminated *č* is found in the word *oččə* (evidently, from **otčə*) ‘vessel made of birch bark’ and may optionally be created as a result of morphological derivation, e.g. *terikəd-jə* ‘married (INTR.1SG)’ > *terikəd-čə* > *terikət-čə* (> *terkəc-čə*).

3.3. Vowel harmony

Vowel harmony operates in two domains, the first foot and the uninflected word, and the rules differ. The first foot is bimoraic, and so has the following structures: (C)VCV, (C)V: or (C)VC (on the foot structure see 3.4.1 and 4.3). Obviously, the harmony is only observed in the (C)VCV type. Note that the distinction between two harmonic domains has not been mentioned in previous work. All previous studies (Jochelson 1905; Collinder 1940; Krejnović 1982; Maslova 2003) describe one type of harmony and extrapolate it onto the word as a whole.

3.3.1. *Harmony in the first foot*

Stems are either harmonically back or harmonically front. The distribution of vowels into harmonic classes according to frontness is shown in (4).

(4)	front	e	ö	u	i
	back	a	o	(u)	i

If the first foot has two vowels, i.e. has the structure (C)VCV, these vowels must share the value for the frontness feature. The front and back vowels do not co-occur within the first foot. The value of the frontness feature of the stem determines the harmonic value of some long vowels in further syllables (3.3.3), as well as the synharmonism (3.3.4).

As follows from (4), *e* and *ö* are front, and *a* and *o* are back. The vowel *u* is harmonically front but phonetically back (for a historical explanation of this fact see 4.1.3). This means that it is compatible with *e* and *ö* (e.g. K *pure*, T *pure* ‘on, outside’, K *čugö* ‘road, trace’), but typically not with *a* and *o*. There are only a few words where *u* and *a* co-occur. Some are recent borrowings, e.g. T *pura:qil* ‘jaeger (Stercorarius)’ ~ Ev. *prakil* ‘kind of bird’ and T *qularqə* ‘seagull’ < Ev. *kular* ‘gull’. Similarly, K *šule-* ‘to stand still’ is a back stem. The vowel *i* occurs both with front and back vowels, as demonstrated in (5a) and (5b), respectively.

(5)a.	kise-	show	K T ijer	apart
	b. ibor	wart	mido-, T mira-	to roam
	čitata:	ground squirrel	K T imol	saddle reindeer
	T cibaya-	to peep (of a mouse)		

Although the latter cases are less frequent, they are by no means uncommon. For a historical explanation of this see 4.1.3.

The vowels of the first foot also share the roundness features, but the roundness harmony only operates on a subset of vowels, namely, on non-front non-high vowels. In other words, the back vowels *o* and *a* cannot co-occur within the first foot, so the structures *CaCo or *CoCa are generally forbidden. In (6a) I present examples of roots with non-front vowels, and in (6b) of roots with front vowels.

(6) a.	K T aña	mouth	b. ögö	fish trap
	K T nono	loop	mere-	to fly

A number of exceptions such as *toyal* ‘scum’ or T *mojarqə* ‘softened (fish)’ are insignificant. They may result from erroneous recording.

Vowels in front stems do not necessarily agree in roundness, e.g. T *köde* ‘man’, T *köne* ‘friend’. Such stems normally have the structure CöCe (but not *CeCö), while the first consonant is the labial *p* or *m*, sometimes also *k* or *š/s*. This suggests that originally the stem was CeCe, but the first vowel was labialized under the influence of the word-initial consonant (see 4.1.2). In other words, the lack of rounding harmony is here secondary.

3.3.2. Harmony of short vowels after the first foot

After the first foot the harmony is mainly restricted to short non-high vowels. All high vowels, most of which are epenthetic (see 3.6), are harmony-neutral. As was mentioned in 3.1.1, the only short non-high vowel after the first foot is *ə*. In roots and derivational suffixes this vowel can harmonize to the vowels of the first foot. Unlike in the first foot, this kind of harmony is optional. Some speakers always pronounce *ə*, while in the pronunciation of other speakers this vowel shows partial harmony in frontness and roundness to the vowel(s) of the first foot in the following manner.

(7)	first foot	harmony	examples
	a	a	[šaqa] or [šaqa] fox
	o	o	[monoyo] or [monoyə] moufflon
	e	e	[terike] or [terikə] wife

As shown in Nikolaeva (1988a), harmony is more likely to occur in the stressed position, while the non-stressed *ə* normally remains phonetically neutral.

This kind of optional assimilation of *ə* has been described as harmony in previous studies on Yukaghir and is the main reason for the transcriptional inconsistencies in the existing sources. The vowel *ə* may be rendered as *a*, *o*, or *e*, even within the same word and when recorded by the same author. This is shown by the following examples.

- (8) a. a:čə ‘domestic reindeer’
 KK a:t'e, at'e, a:c'a, a:c'e; KJ a:če; KD a:če; SD aca, ase-, ace-; SU -ačā; RS -ačā; M áčā, aāčē, aāčā; MO amd'e [rect. atčē]; B aaitsha; ME ointscha; MK áatsche, aatschen-

b. *toukə* ‘dog’

KK *touke*, *tewke-*, *towke-*, *touka*; KJ *toboko*, *toboke*; SD *tobuka*; RS *towēka*; M *towóka*, *tówoka*, *towóko-*, *towóka-*, *tóuka*; MC *tawala* [rect. *tawaka*]; B *tabaka*; ME *tavacke*

c. *unemə* ‘ear’

KK *uneme-*; KJ *uneme*; SD *unume*; SU *unēma*; RS *unama*; MC *inama*; BO *inēmo-*; KL *unemo*; B *oonomma*; ME *unoma*; MK *unjáma*

d. *šoromə* ‘man’

KK *šoromo*; KJ *šoromo*, *romo*; KD *coromo*, *coro*, *romo-*; SD *šoromo*; TD *-soromo*; SU *čoróma*; RS *-šoroma*; M *soróma*, *sorómaga*; KL *šoroma*, *šoromo*; B *toromma*, *-toromma*; ME *torroma*, *-schoromok*; MK *tschrróma*

Such variations do not normally occur with short vowels within the first foot, which are rendered in the same way by all authors, cf.

(9) a. *aŋa* ‘mouth’

KK *aŋa*; KJ *aŋa*; KD *aŋa*; SU *ága*; RS *angá*; MC *aigga* [rect. *angga*], *anga*; MO *aigg* [rect. *angg*]; BO *anáŋga*; KL *angaga*; B *angá*; ME *anga*; MU *anggá*; MK *ánga*

b. *moyo* ‘hat’

KJ *moyo*; SD *mogo*; B *mogo*; MU *mongó*, but ME *mochga*

Inflectional morphemes are normally excluded from the harmony. By inflectional I mean case and number suffixes for nouns and tense/mood and agreement suffixes for verbs, as well as suffixes of non-finite verbs. In contrast, verbal aspectual and voice-like affixes do harmonize. In inflectional morphemes all non-high short vowels are pronounced as *ə* or (stressed) *e* (stressed *a* in Tundra Yukaghir). For example, *moyo-gə* ‘hat (LOC)’ may be pronounced as [*moyoge*] or [*moyogə*] and *aŋa-lə* ‘mouth (INSTR)’ may be pronounced as [*aŋale*] or [*aŋalə*].