The Tungusic Languages

Edited by Alexander Vovin, José Andrés Alonso de la Fuente and Juha Janhunen

Routledge Language Family Series

Routledge

THE TUNGUSIC LANGUAGES

The Tungusic Languages is a survey of Tungusic, a language family which is seriously endangered today, but which at the time of its maximum spread was present all over Northeast Asia.

This volume offers a systematic succession of separate chapters on all the individual Tungusic languages, as well as a number of additional chapters containing contextual information on the language family as a whole, its background and current state, as well as its history of research and documentation. Manchu and its mediaeval ancestor Jurchen are important historical literary languages discussed in this volume, while the other Tungusic languages, around a dozen altogether, have always been spoken by small, local, though in some cases territorially widespread, populations engaged in traditional subsistence activities of the Eurasian taiga and steppe zones and the North Pacific coast.

All contributors to this volume are well-known specialists on their specific topics, and, importantly, all the authors of the chapters dealing with modern languages have personal experience of linguistic field work among Tungusic speakers.

This volume will be informative for scholars and students specialising in the languages and peoples of Northeast Asia, and will also be of interest to those engaged with linguistic typology, cultural anthropology, and ethnic history who wish to obtain information on the Tungusic languages.

Alexander Vovin (†), Directeur d'études, École des hautes études en sciences sociales (EHESS), Paris, France.

José Andrés Alonso de la Fuente, Associate Professor, Institute of Linguistics, Translation Studies and Hungarian Studies, Jagiellonian University, Kraków, Poland.

Juha Janhunen, Professor Emeritus of East Asian Languages and Cultures, University of Helsinki, Finland.

ROUTLEDGE LANGUAGE FAMILY SERIES

Each volume in this series contains an in-depth account of the members of some of the world's most important language families. Written by experts in each language, these accessible accounts provide detailed linguistic analysis and description. The contents are carefully structured to cover the natural system of classification: phonology, morphology, syntax, lexis, semantics, dialectology, and sociolinguistics.

Every volume contains extensive bibliographies for each language, a detailed index and tables, and maps and examples from the languages to demonstrate the linguistic features being described. The consistent format allows comparative study, not only between the languages in each volume, but also across all the volumes in the series.

The Germanic Languages
Edited by Ekkehard Konig and Johan van der Auwera

The Turkic Languages, 2nd Edition
Edited by Éva Csató and Lars Johanson

The Indo-Aryan Languages
Edited by George Cardona and Dhanesh K. Jain

The Austronesian Languages of Asia and Madagascar
Edited by Alexander Adelaar and Nikolaus Himmelmann

The Tai-Kadai Languages
Edited by Anthony V. N. Diller, Jerold A. Edmondson and Yongxian Luo

The Iranian Languages
Edited by Gernot Windfuhr

The Khoesan Languages
Edited by Rainer Vossen

The Munda Languages
Edited by Gregory D.S. Anderson

The Celtic Languages, 2nd Edition
Edited by Martin J. Ball and Nicole Müller

The Sino-Tibetan Languages, 2nd Edition
Edited by Graham Thurgood and Randy J. Lapolla

The Indo-European Languages, 2nd Edition
Edited by Mate Kapović

The Mayan Languages
Edited by Judith L. Aissen, Nora C. England, and Roberto Zavala Maldonado

The Languages of Japan and Korea
Edited by Nicolas Tranter

The Mongolic Languages
Edited by Juha Janhunen

The Oceanic Languages
Edited by John Lynch, Malcolm Ross and Terry Crowley

Language Isolates
Edited by Lyle Campbell

The Bantu Languages, 2nd Edition
Edited by Mark Van de Velde, Koen Bostoen, Derek Nurse, and Gérard Philippson

The Semitic Languages, 2nd Edition
Edited by John Huehnergard and Na'ama Pat-El

The Dravidian Languages, 2nd Edition
Edited by Sandford B. Steever

The Uralic Languages, 2nd Edition
Edited by Daniel Abondolo and Riitta-Liisa Valijärvi

The Tungusic Languages
Edited by Alexander Vovin, José Andrés Alonso de la Fuente, and Juha Janhunen

Other titles in the series can be found at https://www.routledge.com/Routledge-Language-Family-Series/book-series/SE0091

THE TUNGUSIC LANGUAGES

Edited by Alexander Vovin,
José Andrés Alonso de la Fuente,
and Juha Janhunen

Routledge
Taylor & Francis Group
LONDON AND NEW YORK

First published 2024
by Routledge
4 Park Square, Milton Park, Abingdon, Oxon OX14 4RN

and by Routledge
605 Third Avenue, New York, NY 10158

Routledge is an imprint of the Taylor & Francis Group, an informa business

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library

Library of Congress Cataloging-in-Publication Data
Names: Vovin, Alexander, editor. | Alonso de la Fuente, José Andrés, editor. | Janhunen, Juha, 1952– editor.
Title: The Tungusic languages / edited by Alexander Vovin, José Andrés Alonso de la Fuente, and Juha Janhunen.
Description: Abingdon, Oxon ; New York, NY : Routledge, 2023. | Series: Routledge language family | Includes bibliographical references and index.
Identifiers: LCCN 2023004787 (print) | LCCN 2023004788 (ebook) | ISBN 9781138845039 (hardback) | ISBN 9781032463940 (paperback) | ISBN 9781315728391 (ebook)
Subjects: LCSH: Tungus-Manchu languages. | LCGFT: Essays.
Classification: LCC PL450 .T86 2023 (print) | LCC PL450 (ebook) | DDC 494/.1—dc23/eng/20230501
LC record available at https://lccn.loc.gov/2023004787
LC ebook record available at https://lccn.loc.gov/2023004788

ISBN: 978-1-138-84503-9 (hbk)
ISBN: 978-1-032-46394-0 (pbk)
ISBN: 978-1-315-72839-1 (ebk)

DOI: 10.4324/9781315728391

Typeset in Times New Roman
by Apex CoVantage, LLC

CONTENTS

TABLES

CONTRIBUTORS

José Andrés Alonso de la Fuente
Associate Professor, Institute of Oriental Studies, Faculty of Philology
Jagellonian University, Kraków, Poland

Sangyub Baek
Associate Professor, Center for International Relations
Muroran Institute of Technology, Muroran, Japan

Gao Wa
Postdoctoral Researcher
Central University for Nationalities, Peking, PR China

Juha Janhunen
Professor Emeritus of East Asian Languages and Cultures
Department of Languages, University of Helsinki, Helsinki, Finland

Daniel Kane (1948–2021) (1948–2021)
Professor of Chinese, Department of International Studies
Macquarie University, Sydney, Australia

Shinjiro Kazama
Professor of Linguistics, Graduate School of Global Studies
Tokyo University of Foreign Studies, Fuchu, Japan

Andrej Malchukov
Außerplanmäßiger Professor of Linguistics, Department of English and Linguistics
Johannes Gutenberg University, Mainz, Germany

Nadezhda Mamontova
Newton International Fellow, School of Geography, Earth and Environmental Sciences,
University of Birmingham, UK
Visiting scholar, Cultural History and European and World History, University of Turku,
Finland

Marc Hideo Miyake
Independent Researcher
‘Aiea, Hawai‘i, USA

Sofia Oskolskaya
Researcher, Department of the Languages of Russia, Institute for Linguistic Studies
Russian Academy of Sciences, St. Petersburg, Russia

Elena Perekhvalskaya
Researcher, Department of the Languages of Russia, Institute for Linguistic Studies
Institute for Linguistic Studies, RAS, St. Petersburg, Russia

Toshiro Tsumagari (1951–2020)
Professor Emeritus, Hokkaido University, Sapporo
Director, Hokkaido Museum of Northern Peoples, Abashiri, Japan

Alexander Vovin (1961–2022)
Directeur d'études, Centre nationale de la recherche scientifique (CNRS)
École des hautes études en sciences sociales (EHESS), Paris, France

Lindsay J. Whaley
Professor of Linguistics and Classics, Department of Linguistics
Dartmouth College, Hanover, New Hampshire, USA

Yoshiko Yamada
Associate Professor, College of Liberal Arts
Muroran Institute of Technology, Muroran, Japan

Veronika Zikmundová
Deputy Head, Institute of Asian Studies, Faculty of Arts
Charles University, Prague, Czechia

PREFACE AND ACKNOWLEDGEMENTS

This volume is a survey of Tungusic, a language family which is seriously endangered today, but which at the time of its maximum spread just a couple of centuries ago was present all over Northeast Asia, with the speakers of one particular Tungusic language, Manchu, even ruling over several neighbouring regions, including, most importantly, China. Manchu and its mediaeval ancestor Jurchen are also important historical literary languages, while the other Tungusic languages, around a dozen altogether, have always been spoken by small local, though in some cases territorially widespread, populations engaged in traditional subsistence activities of the Eurasian taiga and steppe zones and the North Pacific coast.

Following the example of other volumes in the Routledge series on language families, including, in particular, those on Turkic (1998, 2021) and Mongolic (2003), this volume offers a systematic succession of separate chapters on all the individual Tungusic languages, as well as a number of additional chapters containing contextual information on the language family as a whole, its background and current state, as well as its history of research and documentation. All contributors to this volume are well-known specialists on their specific topics, and, importantly, all the authors of the chapters dealing with modern languages have personal experience of linguistic field work among Tungusic speakers.

Technically, all chapters on the individual languages follow a roughly uniform template, starting with an ethnolinguistic introduction and a history of documentation, and proceeding through phonology, morphophonology, phonotactics, as well as derivational and inflectional morphology to a brief description of syntax, lexicon, and external contacts, as well as a list of references. The general emphasis in all chapters is on the synchronic structural description of the target language against a diachronic and comparative background. Since this is the first comprehensive handbook of Tungusic in English, it has been an important task to create and establish a generally acceptable standard of notation and terminology. In this respect, the chapters have been unified during the editorial process.

In spite of the unified format of the descriptive chapters, the authors have had the freedom to focus on the topics they consider important. For instance, although the general focus of most chapters is on phonology and morphology, some chapters offer relatively more information on syntax. The ordering of the data in the individual chapters also varies to some extent, reflecting the preferences of the authors and the specific characteristics of each individual language. The amount of diachronic and comparative information likewise varies, though such information has also been added to some chapters during the editorial process as far as required by the general framework. The editors are convinced that diachrony is an essential clue to understanding synchrony, and it is a particularly important tool when the facts concerning an entire language family are to be summarized.

This book has been very long in the making. Originally commissioned to Alexander (Sasha) Vovin soon after 2003, a first attempt to collect manuscripts from authors was completed around 2010. However, this attempt did not result in publication, and most of the texts written then were ultimately published elsewhere, or have been rewritten and included in the present volume in the form of updated versions. Due to his involvement in other projects, Sasha Vovin subsequently invited José Andrés Alonso de la Fuente to be a coeditor and the active contact person coordinating the correspondence with the authors. However, in late 2020, the two initial editors asked Juha Janhunen to complete the actual editing work along lines similar to the volume on Mongolic, which he had edited earlier.

Due to the prolonged editorial process, many changes have taken place in the original plan, including the choice of the topics, the ordering of the chapters, and even the selection of the authors. Most importantly, and most sadly, three of the original authors have passed away over the years. Most recently, Sasha Vovin (1961–2022) himself died only several months before the completion of the project, while Tsumagari Toshiro (1951–2020) and Daniel Kane (1948–2021) also passed away without having seen the final versions of their texts. However, the chapters of both Kane and Tsumagari have been completed by their coauthors Marc Hideo Miyake and Yamada Yoshiko, respectively. As far as possible, all chapters have been brought up to date by the authors and the editors.

As editors we express our thanks to all authors for their valuable contributions and cooperation, as well as for their patience. Many other colleagues have also provided us comments and answered our questions in the process of our work, including, but not limited to Uwe Bläsing (Leiden), Tom Eriksson (Vilnius), Stefan Georg (Bonn), Michael Knüppel (Göttingen), Li Baowen (Peking), Alfred F. Majewicz (Poznań), Matti Miestamo (Helsinki), I. N. Novgorodov (Yakutsk), Mehmet Ölmez (Istanbul), William Rozycki (Bloomington), Fresco Sam-Sin (Leiden), Marek Stachowski (Cracow), Giovanni Stary (Venice), and Jerzy Tulisow (Warsaw). Our thanks go also to the Routledge staff Andrea Hartill and Iola Ashby for their patience, understanding and smooth cooperation. The editorial input of Juha Janhunen has been supported by the Laboratory Program for Korean Studies through the Ministry of Education of the Republic of Korea and Korean Studies Promotion Service of the Academy of Korean Studies (AKS-2016-LAB-2250004).

10 December 2022

José Andrés Alonso de la Fuente (Cracow, Poland)

Juha Janhunen (Helsinki, Finland)

TECHNICAL NOTES

Material from spoken languages is transcribed in the present volume in a unified system based on the phonemic application of the basic Roman letters, complemented by a selection of special symbols and diacritically modified letters. The general principle is that one letter corresponds to one phonemic segment. Distinctive long vowels, diphthongs, and vowel sequences, as well as consonant geminates, are written as sequences of two letters. Otherwise, sequential notation is used only in a few special cases which are open to alternative phonemic analyses (such as the different phonation types of vowels in Udihe). The correspondence of the letters to their phonetic values is approximate and varies from language to language. For a more exact description of the underlying sounds in each given language, the International Phonetic Alphabet (IPA) is applied when necessary.

Most Tungusic languages have a relatively simple consonant system organized in terms of four places of articulation: labial, dental, palatal, and velar, and comprising up to four nasals written as *m n ñ ŋ*, two sets of stops phonemically characterized as "strong" vs. "weak" and written as *p t c k vs. b d j g*, two fricatives written as *s* (dental sibilant) and *x* (velar to laryngeal), two glides written as *w* (labial, in some idioms with a dentilabial realization) and *y* (palatal), and two liquids written as *l* (lateral) and *r* (vibrant/trill). The palatal stops *c j* are in most languages pronounced as affricates (in some idioms also with dental values). Additional consonants present only in a few languages include *f* (labial fricative with both bilabial and dentilabial realizations), *š* (distinctive palatal sibilant), *q ɢ* (distinctive postvelar or uvular stops), as well as *h* (distinctive postvelar to laryngeal fricative or spirant).

The vowel systems show more variation. In the simplest form, as attested in several Tungusic idioms, the system comprises two rounded back vowels written as *o u* (mid-high vs. high), two unrounded central or back vowels written as *a e* (low vs. mid-high), and one unrounded front vowel written as *i* (high). It is important to note that the letter *e* is here used for a central or back vowel quality, corresponding to the phonetic definition of "schwa" [ə]. The corresponding front vowel [e] is written, when necessary, as *é* or *iě*. Other vowels, occurring as distinctive segments in several languages include *ʊ ɪ* (lowered counterparts of *u i*) and *ɵ* (a higher or centralized counterpart of *o*), as well as *ä ö ü* (palatal counterparts of *a o u*). In reconstructions, the letter *ï* (the velar counterpart of *i*) is also used.

As far as written languages using a non-Roman script are concerned, established transcription or transliteration systems are used for Written Manchu (Möllendorff-Norman), Written Mongol (Balk-Janhunen), Mandarin Chinese (Pinyin, mainly without tones), Japanese (Hepburn), and Korean (either McCune-Reischauer or Revised Romanization, depending on the author), with a consideration of individual preferences in the case of personal names. For the Manchu script, see also the separate chart in this volume. For Russian, a mixed system (with *c ch sh shh zh kh* for ц ч ш щ ж х and *è e ë ya yu y i ï ’ ”* for э е ё я ю ы и й ии ь ъ) is used. All bibliographical references are, however, provided

in the original script, as are the Chinese characters (mainly in unsimplified form) for relevant terms and names in the chapter texts. Note that many modern Tungusic languages have one or several written standards based on either Cyrillic or Roman letters. Any linguistic data from these languages is, however, quoted in the unified notation as used in this volume.

Language data, including reconstructions, Romanizations from Chinese, Japanese, Korean, and Russian, as well as tentative "readings" from non-alphabetic scripts (as in the case of Jurchen written in Chinese characters) are quoted in italics, while Romanized Manchu and Mongol are quoted in boldface. Reconstructions are marked by the asterisk (*), and "readings" are marked by the dagger (†). In some cases, the radical symbol (√) is used to indicate lexical roots. Since the transcription of the spoken languages is intended to be basically phonemic, phonemic slashes (/ /) are used only in specific cases when required by the discourse. Phonetic data in IPA are given in square brackets ([ ]), while angle brackets (⟨ ⟩) indicate graphic segments (letters) and sequences. The less vs. greater than signs (< >) stand for the direction of diachronic evolution, and the arrows (← →) show the direction of borrowing. The hyphen (-) indicates morpheme boundaries, the equal sign (=) boundaries of clitics, the plus sign (+) word boundaries in compound words, and the ampersand (&) the boundary between reduplicated elements. Glosses are indicated by the single quotes (‘ ’). In some cases connective segments or etymological morpheme boundaries are indicated by the period mark (.), while the colon (:) surrounded by a space on both sides indicates paradigmatic relationships.

The ethnic nomenclature has been unified throughout the volume, as explained in Chapter 1. For grammatical terminology, the principle has been to use, in the first place, terms already established in Tungusic studies, but since many features lack an established term, especially in English, this volume also uses a number of innovative terms, especially for features for which the Tungusic languages themselves exhibit no uniform picture (as in the systems of converbs). Some new terms have also been adopted from mainstream linguistics. All terms are included in the List of abbreviations. For historical datings, the abbreviations BZ (Before Zero = BC) and AZ (After Zero = AD) are used.

ABBREVIATIONS

⟨1⟩	first person
⟨2⟩	second person
⟨3⟩	third person
⟨ABL⟩	ablative
⟨ACC⟩	accusative
⟨ADD⟩	additive
⟨ADJ⟩	adjectival
⟨ADN⟩	adnominal
⟨ADV⟩	adversative
⟨AFF⟩	affirmative
⟨AG⟩	agentive
⟨AL⟩	alienable
⟨ALT⟩	alternative
⟨AND⟩	andative
⟨ANT⟩	anterior
⟨ANTICAUS⟩	anticausative
⟨AOR⟩	aorist
⟨APPR⟩	apprehensive
⟨ASSOC⟩	associative
⟨AUGM⟩	augmentative
⟨AUX⟩	auxiliary
⟨CAUS⟩	causative
⟨CL⟩	clitic
⟨COLL⟩	collective
⟨COM⟩	comitative
⟨COMP⟩	comparative
⟨COMPL⟩	completive
⟨CONC⟩	concessive
⟨COND⟩	conditional
⟨CONF⟩	confirmative
⟨CONJ⟩	conjunctive
⟨CONN⟩	connective
⟨CONNEG⟩	connegative
⟨CONT⟩	continuative
⟨CONTR⟩	contrastive
⟨COOP⟩	cooperative
⟨COP⟩	copula
⟨CORR⟩	corrogative
⟨CSEC⟩	consecutive
⟨CTEMP⟩	contemporal
⟨CTXT⟩	contextual
⟨CV⟩	converb
⟨CX⟩	case suffix
⟨DAT⟩	dative
⟨DEB⟩	debitive
⟨DEL⟩	delative
⟨DER⟩	derivative
⟨DES⟩	designative
⟨DESID⟩	desiderative
⟨DEST⟩	destinative
⟨DIM⟩	diminutive
⟨DIR⟩	directive
⟨DISC⟩	discursive
⟨DISJ⟩	disjunctive
⟨DIST⟩	distal
⟨DISTR⟩	distributive
⟨DN⟩	denominal
⟨DV⟩	deverbal
⟨DX⟩	derivational suffix
⟨E⟩	epenthetic
⟨ELAT⟩	elative
⟨EMPH⟩	emphatic
⟨EQU⟩	equative
⟨EVID⟩	evidential
⟨EXCL⟩	exclusive
⟨EXIST⟩	existential
⟨EXPR⟩	expressive
⟨EXT⟩	extension
⟨FEM⟩	feminine
⟨FIN⟩	finite
⟨FOC⟩	focus
⟨FUT⟩	futur(itiv)e
⟨HAB⟩	habitive
⟨HORT⟩	hortative
⟨HYP⟩	hypothetical
⟨GEN⟩	genitive
⟨IMMED⟩	immediative
⟨IMP⟩	imperative

⟨IMPERS⟩	impersonal
⟨IMPRF⟩	imperfective
⟨INCH⟩	inchoative
⟨INCL⟩	inclusive
⟨INDEF⟩	indefinite
⟨INFIN⟩	infinitive
⟨INSTR⟩	instrumental
⟨INTENS⟩	intensive
⟨INTENT⟩	intentive
⟨INTERJ⟩	interjection
⟨INTERR⟩	interrogative
⟨ITER⟩	iterative
⟨ITR⟩	intransitive
⟨LAT⟩	lative
⟨LIM⟩	limitative
⟨LOC⟩	locative
⟨MASC⟩	masculine
⟨MED⟩	medial
⟨MIR⟩	mirative
⟨MOD⟩	modal
⟨MODER⟩	moderative
⟨MOM⟩	momentative
⟨MULT⟩	multiplicative
⟨MULTID⟩	multidirectional
⟨N⟩	noun
⟨NEC⟩	necessitative
⟨NEG⟩	negative
⟨NMLZ⟩	nominalization
⟨NONDUM⟩	nondumitive
⟨NX⟩	number suffix
⟨OBL⟩	oblique
⟨OBV⟩	obviative
⟨OCX⟩	oblique coaffix
⟨ONOM⟩	onomatopoetic
⟨OPT⟩	optative
⟨ORD⟩	ordinal
⟨P⟩	person
⟨PART⟩	partitive
⟨PASS⟩	passive
⟨PEJ⟩	pejorative
⟨PERM⟩	permissive
⟨PERS⟩	personal
⟨PL⟩	plural
⟨PLUPRF⟩	pluperfect
⟨PLUR⟩	pluritative
⟨POSS⟩	possessive
⟨PRET⟩	pretentive
⟨PREV⟩	preventive
⟨PRF⟩	perfective
⟨PROB⟩	probabilitive
⟨PROGR⟩	progressive
⟨PROL⟩	prolative
⟨PROPR⟩	proprietive
⟨PROSP⟩	prospective
⟨PROX⟩	proximal
⟨PRS⟩	present
⟨PST⟩	past
⟨PTCP⟩	participle
⟨PTCL⟩	particle
⟨PURP⟩	purposive
⟨PX⟩	possessive suffix
⟨RECIPR⟩	reciprocal
⟨REDUPL⟩	reduplication
⟨REF⟩	reference
⟨REFL⟩	reflexive
⟨REM⟩	remote
⟨RES⟩	resultative
⟨RESTR⟩	restrictive
⟨REV⟩	reversive
⟨RX⟩	reflexive suffix
⟨SEM⟩	semelfactive
⟨SEP⟩	separative
⟨SG⟩	singular
⟨SIM⟩	similative
⟨SING⟩	singulative
⟨STAT⟩	stative
⟨SUP⟩	supine
⟨SUPERL⟩	superlative
⟨TERM⟩	terminative
⟨TOP⟩	topic
⟨TR⟩	transitive
⟨TRANSL⟩	translative
⟨V⟩	verb
⟨VBLZ⟩	verbalization
⟨VOC⟩	vocative
⟨VOL⟩	voluntative
⟨VOLIT⟩	volitive
⟨VX⟩	personal ending

CHART OF THE MANCHU SCRIPT

In this volume, the letters of the Manchu alphabet are Romanized according to the system of Jerry Norman (e.g. Norman 2013: xv–xxiv), which, in its turn, is a slightly modified version of the system introduced by Paul Georg von Möllendorff (1892). There have been many proposals of Romanization in the past, but all of them involve problems in the details, not only because of the constant confusion between the principles of transcription vs. transliteration, but also because of the presence of special conventions used when writing foreign words in the Manchu script (cf. Osterkamp 2016).

The chart below covers the letters used to represent the sounds of the inherited Tungusic vocabulary of Manchu but only a basic set of letters devised for the transcription of Chinese and other loanwords. Apart from the system of Norman, the chart includes the Romanizations used in Hauer (2007) and Hu Zengyi (ed.) (1994: 894–905), complemented by the Cyrillic transcription of Zakharov (1875: 3–14). For more comprehensive comparative tables, cf. Stary (2000: xvi–xvii) and, especially, Lie (1972: 72–77).

	Norman	Hauer	Hu Zengyi	Zakharov
ᡳ	**a**	a	a	а
ᡦ	**b**	b	b	б
ᠴ	**c**	c	ch	ч (ц)
ᠴᡳ	**cy**	c’	chy	чы
ᡩ ᡯ	**d**	d	d	д
ᡝ	**e**	e	e	э
ᡶ ᡨ	**f**	f	f	ф
ᡤᡳ	**g**	g	g	г, г̄
ᡤᡤ	**g’**	g’	gg	г̄
ᡥᡳ	**h**	h	h	х, х̄
ᡥᡥ	**h’**	h’	hh	(-)
ᡳ	**i**	i	i	и, і, й

	Norman	Hauer	Hu Zengyi	Zakharov
ᡤ	**j**	j	zh	чж
ᡤᡳ	**jy**	‘j	zhy	чжи
ᡨ	**ž**	ǰ	rr	ж
ᡴ ᠺ	**k**	k	k	к, к̄
ᡵ	**k’**	k’	kk	к̄
ᡶ	**l**	l	l	л
ᡷ	**m**	m	m	м
ᡶᡳ	**n**	n	n	н
ᡮ	**o**	o	o	о
ᡯ	**p**	p	p	п
ᡰ	**r**	r	r	р
ᡧ	**s**	s	s, s’	с
ᡧᡳ	**sy**	s	sy	сы
ᡧᡠ	**š**	š	sh	ш
ᡨ ᡨᡳ	**t**	t	t	т
ᡮᡠ	**u**	u	u	у
ᡮᡠᡳ	**ū**	û, ô	uu	у, ӯ
ᡰᡳ	**w**	w	w	в
ᡳ	**y**	y	y	(-)
ᡦ	**dz**	z	z	цз
ᡦᡳ	**ts**	z’	c	ц
ᡢ ᡢᡳ	**ng**	ng	ng	-нг, -н-

Note that in the system of Hu Zengyi ⟨s’⟩ is used to disambiguate the sequence ⟨s⟩ *cum* ⟨h⟩, i.e., [s.x ~ s.χ] from ⟨sh⟩ [ʂ], as, for example, in ⟨tas’ha⟩ ‘tiger’ (for **tasha**, Zakharov тасха). In Zaxarov’s system, ⟨я ѣ іō ю⟩ mark sequences with a palatal element (**iya iye iyo iyū**). The two lines on some consonant letters in the chart indicate combinations with a following “male” or “female” vowel (**d k t**), or occurrence in initial vs. medial vs. final position (**f ng**). In fact, due to the history and typology of the script, several other letters have also separate graphic forms for the initial, medial, and final, as well as, in some cases, absolute position, of which the above chart includes only the initial forms, with the exception of the letter **sy**, which, for technical reasons, is given in its complex initial-final form.

REFERENCES

Corff, Oliver & Kyoko Maezono & Wolfgang Lipp & Dorjpalam Dorj & Göröochin Gerelmaa & Aysima Mirsultan & Réka Stüber & Byambajav Töwshintögs & Xieyan Li (eds.) (2013) *Auf kaiserlichen Befehl erstelltes Wörterbuch des Manjurischen in fünf Sprachen—"Fünfsprachenspiegel"*, vols. 1–2 + indices 1–5, Wiesbaden: Harrassowitz.

Hauer, Erich (2007) *Handwörterbuch der Mandschusprache*, 2., durchgesehene und erweiterte Auflage herausgegeben von Oliver Corff, Wiesbaden: Harrassowitz Verlag.

Hu, Zengyi 胡增益 (ed.) (1994)『新满汉大词典』[A Comprehensive Manchu-Chinese Dictionary], 烏魯木齊 [Urumqi]: 新疆人民出版社.

Lie, Hiu (1972) *Die Mandschu Sprachkunde in Korea*, Indiana University Uralic and Altaic Series 114, Bloomington, Indiana: Indiana University, Research Institute for Inner Asian Studies.

Möllendorff, P[aul] G[eorg] von (1892) *A Manchu Grammar with Analyzed Texts*, Shanghai: American Presbyterian Mission Press.

Norman, Jerry (2013) with the assistance of Keith Dede and David Prager Branner, *A Comprehensive Manchu-English Dictionary*, Cambridge (Massachusetts) and London: Harvard University Asia Center.

Osterkamp, Sven (2016) Review of: Corff & al (2013), *Asiatische Studien—Études Asiatiques* 70 (3): 942–965, Berlin.

Stary, Giovanni (2000) *A Dictionary of Manchu Names: A Name Index to the Manchu Version of the "Complete Genealogies of the Manchu Clans and Families of the Eight Banners"*, Aetas Manjurica 8, Wiesbaden: Otto Harrassowitz.

Zakhrov, Ivan [Иванъ Захаровъ] (1875) *Полный маньчжурско-русскій словарь* [A complete Manchu-Russian dictionary], Санктпетербургъ [St. Petersburg]: Типографія Императорской Академіи Наукъ.

LANGUAGE MAP

This map shows the locations and maximal territories of the Tungusic languages at a time level corresponding broadly to the ethnographic present of the 17th to 19th centuries. The languages are indicated by numbers corresponding to the chapters in this volume. For more detailed maps, indicating individual localities and/or clans and tribal-level groups, see, in particular, Dolgikh (1960 map), Vasilevich (1958 map), and Service (1978 map).

7 Siberian Ewenki	10 Neghidal	13 Udihe	16 Uilta
8 Orochen	11 Ewen	14 Nanai	17 Manchu
9 Solon	12 Oroch	15 Ulcha	18 Sibe

Note that the western expansion of Ewenki towards the Ob basin and the northern and northeastern expansion of Ewen towards the Anadyr' basin and the Arctic coast was still in progress during this period, while the expansion of Yakut was gradually hollowing out the territories of both Ewenki and Ewen in the Lena basin, which also marked the approximate mutual boundary of Ewenki and Ewen in the north. Solon underwent translocations from the Middle Amur basin to the Nonni basin in the late 17th century and further to the Hailar basin in the early 18th century, while still another translocation took it, together with Sibe, to Jungaria in 1764. The presence of Ewen on central Kamchatka and Ewenki on central Sakhalin dates back only to the 19th century. In the course of the 20th century, the territories of all Tungusic languages have shrunk to a network of points marking the settlements where the remaining speakers live. As a case of exception, broader expanses of space are still occupied locally by the northernmost reindeer herding Ewenki and Ewen speakers, but even in these cases the bulk of the population lives in multiethnic fixed settlements.

REFERENCES

Dolgikh, B. O. [Б. О. Долгих] (1960) *Родовой и племенной состав народов Сибири в XVII в.* [Clan and tribe composition of the peoples of Siberia in the 17th c.], Труды Института этнографии им. Н. Н. Милухо-Маклая, Новая серия, том LV, Москва [Moscow]: Издательство Академии Наук СССР.

Service, Robert T. (1978) 'Notes on a map of the Tungusic peoples and their dialects', in: Gerhard Doerfer & Michael Weiers (eds.), *Beiträge zur nordasiatischen Kulturgeschichte*, Tungusica 1: 52–65, Wiesbaden: Otto Harrassowitz.

Vasilevich, G. M. [Г. М. Василевич] (1958) *Эвенкийско-русский словарь* [Ewenki-Russian dictionary], Москва [Moscow]: Государственное издательство иностранных и национальных словарей.

CHAPTER 1

TUNGUSIC AS A LANGUAGE FAMILY

Juha Janhunen

Tungusic is a medium-sized language-family with some 15 different languages distributed over a vast territory in Northeast Asia, including both Siberia and Manchuria, and today divided between the Russian Federation and the People's Republic of China, giving it a transnational status. In terms of territorial spread Tungusic is, however, a "large" language family, comparable with the other "large" language families of northern and central Eurasia, such as Mongolic, Turkic, and Uralic. In terms of demographic weight, again, Tungusic is a "small" language family, for the total number of people speaking Tungusic languages today is hardly more than 30,000, probably even less. As a language family, Tungusic is, in fact, seriously endangered, and many of the individual Tungusic languages are on the verge of extinction, with some recently documented varieties already extinct.

The political status of Tungusic has, however, varied in the course of time, and the presence of both Russia and China in Northeast Asia is of a relatively recent date. Most importantly, one particular Tungusic language, Manchu, has a history the like of which few other languages can show, for it was not only the regionally dominant language of Manchuria, spoken by a considerable population, but also the official first language of the Manchu Empire of the Qing (1616–1911), which comprised not only Manchuria, but also all the neighbouring regions, including Mongolia, Tibet, East Turkestan, and China. Manchu enjoyed a brief period of renewed symbolic status during the short-lived empire of Manchukuo (1932–1945), but subsequently it has been degraded to just one of the 54 "minority languages" of the People's Republic, sharing the status of the other Tungusic languages, which have always been spoken by smaller and politically less influential populations. On the Russian side, the speakers of the Tungusic languages are conventionally included in the category of the "Small Peoples of the Far North" (*narody Krainego Severa, Sibiri i Dal'nego Vostoka*).

As far as their traditional modes of subsistence are concerned, the Tungusic speakers represent several regionally diversified adaptations, comprising the settled small-scale agriculturalists of central Manchuria, the settled or semisettled fishing and hunting populations of the Amur basin in northern Manchuria, the nomadic hunters and fishermen of the Ussuri Taiga in eastern Manchuria, the Mongolian-type pastoral nomads of Transbaikalia, and the nomadic reindeer herders living dispersed over most parts of Siberia, with small local communities also on the island of Sakhalin and the peninsula of Kamchatka. A diaspora group of Manchu speakers, officially known as the Sibe, live since the 18th

DOI: 10.4324/9781315728391-1

century in Jungaria, northern Sinkiang (*Xinjiang*), at the modern border towards Kazakhstan, where they share the lifestyle of the other local populations.

The difference in the economic adaptations correlates with differences in material and social culture, as well as with the genetic composition of the different Tungusic speaking populations. Obviously, many of the local groups today or historically speaking Tungusic languages, represent traces of other ethnic groups which were secondarily, and in some cases very recently, absorbed into the Tungusic speaking sphere. There are, however, several cultural features that are shared by all Tungusic speaking populations. Most importantly, all Tungusic speakers are traditionally adherents of the North Asian type of shamanism, a complex world view which may or may not be classified as a "religion". Not surprisingly, the term *shaman* was also transmitted to international usage from, or via, the Tungusic languages. By comparison, other major religions, including Buddhism, Christianity, and Islam have gained little foothold among Tungusic speakers, though officially Tibetan Buddhism was one of the favoured religions of the Manchu Empire, and Russian Orthodox Christianity has been introduced to most Tungusic speaking groups in Siberia.

ETHNIC NOMENCLATURE

The names used to refer to the individual Tungusic languages are in most cases based on the corresponding ethnonyms used either by their speakers or by their neighbours. The ethnic nomenclature is, however, a complicated issue and involves many sources of confusion. The English term "Tungusic", as used here, is itself a relatively new name for the language family, though analogous names have been current since, at least, several decades, in German (*tungusische Sprachen*) and French (*langues toungouses*). The standard term in Russian usage is still today "Tungus-Manchu" (*tunguso-man'chzhurskie yazyki*), while in English the form "Manchu-Tungus" has occasionally been used. Such double names suggest a basic dichotomy of the language family into "Manchu" and "Tungus", a probably premature conclusion which, in any case, remains to be confirmed. The term "Tungusic" has obvious advantages, one of which is its formal parallelism with "Mongolic" and "Turkic".

It happens that the ethnonyms "Tungus" and "Manchu" are also both of a relatively recent origin. "Manchu" (**manju**) was the term consciously introduced, and perhaps artificially coined, by the Manchu ruling elite around the time of their dynastic rise (1635), until which time Manchu speakers had identified themselves as "Jurchen" (*jušen*). "Tungus", on the other hand, was the term adopted by the Russians in the late 16th century from the local Samoyedic speakers in the Lower Yenisei region in reference to the Siberian "Tungus", today known as Ewenki and Ewen. In both Russian and European usage, the term "Tungus" has later been applied also to the other Tungusic speaking peoples, with the exception of the Manchu, who are normally separated from the "Tungus", though not because of linguistic considerations, but because of their distinct demographic, cultural, and political status.

With the advance of the Russians towards the Pacific coast, starting in the mid 17th century, and especially after the acquisition by Russia of the areas north of the Amur (1858) and east of the Ussuri (1860), specific names came to be applied to the Tungusic speaking groups in this region, subsequently known as the Russian Far East. The names were typically adopted from the immediately neighbouring groups, both Tungusic and non-Tungusic. For some groups, several parallel names were applied. However, after

the introduction of Soviet rule, most of these "old" names were regarded as politically incorrect and were replaced by "new" names, often based on the endonyms used by the indigenous groups themselves, or simply on the appellative nouns meaning 'man, person' in their languages. For this reason, most Tungusic peoples of Siberia and the Russian Far East are known by two names, one of which is "old", while the other one is "new". Both names have been used, and are still occasionally used in western sources on Tungusic, though in most cases the "new" names have gradually gained ground also in international Tungusic studies.

Further variation is caused by the ways how the "new" names are handled in Russian and other languages, depending on both morphological and orthographical issues, including problems concerned with Romanization. The Russian terms have normally three different stem forms, one used in the masculine singular (referring to a male member of the ethnic group) and the generic plural form, the second in the feminine singular form (referring to a female member of the ethnic group), and the third in the corresponding adjectival form. In some cases, there is variation in how the basic masculine singular form is constructed. The correspondence with the actual native shapes of the ethnonyms is often approximate, and the Russian term is typically based on only one dialectal variant of the indigenous ethnonym.

On the Chinese side, the ethnonyms used for the Tungusic speaking groups are partly based on adaptations of the traditional Manchu usage, but the orthographic rendering of the Chinese terms in Latin letters varies depending on the system of Romanization, and also on their degree of adaptation to the rules of Chinese syllable structure. Some ethnonyms are used on both sides of the Sino-Russian border, while in other cases different names are used for a single ethnic group on the two sides of the border. To avoid additional confusion, the names of the Tungusic languages in the present volume are based on a standard created with a parallel consideration of both the Russian and the Chinese name forms (Table 1.1). In some cases, it has been judged more appropriate to use a form closer to the actual indigenous ethnonym, rather than a direct adaption of the Russian or the Chinese name form.

TABLE 1.1 NAMES OF THE TUNGUSIC LANGUAGES

1	2	3	4	5	6	7	8
Ewen	*èvén*	*èvény*	*èvénka*	*èvénskii*		Lamut	eve
Ewenki	*èvénk*	*èvénki*	*èvenkíika*	*èvenkíiskii*	*Ewenke*	Tungus	evn
Nanai	*nanáec*	*nanáicy*	*nanáika*	*nanáiskii*	*Hezhe/n*	Gold	gld
Neghidal	*negidálec*	*negidál'cy*	*negidálka*	*negidal´skii*			neg
Oroch	*óroch*	*órochi*	*órochka*	*órochskii*			oac
Orochen	*orochón*	*orochóny*	*orochónka*	*orochónskii*	*Elunchun*		orh
Solon	*solón*	*solóny*	*solónka*	*solónskii*	*Ewenke*		eve
Udihe	*udègéec*	*udègéicy*	*udègéika*	*udègéiskii*		Taz	ude
Ulcha	*ul'ch*	*úl'chi*	*ul'chánka*	*úl'chskii*		Mangun	ulc
Uilta	*uíl'ta*	*uíl'ta*	*uil'tínka*	*uil'tínskii*		Orok	oaa

Name forms: 1. Recommended standardized name, 2. Russian masculine singular, 3. Russian plural, 4. Russian feminine singular, 5. Russian adjectival form (masculine singular), 6. Chinese name (Pinyin), 7. "old" name, 8. ISO 639–3 language code.

It may be pointed out that many of the Russian ethnonyms, both "old" and "new", are artificial creations and have only a diffuse and arbitrary connection with the endonyms and exonyms of the ethnic groups concerned. Similar issues are observed on the Chinese side. Also, as may be seen, the ISO codes for the Tungusic languages are randomly based on either the "old" or the "new" names, and some languages, like Solon and (Siberian) Ewenki, are lumped together under one single code. Confusion is also caused by the similarity of several ethnonyms, notably Ewenki and Ewen, Ulcha and Uilta, as well as Orochen, Oroch, and Orok. A practical consequence of the ethnonymic similarities is that census data, and the accompanying data concerning the numbers of speakers, are not always reliable.

In addition to the above-listed Tungusic languages, whose speakers are recognized as distinct ethnic groups, there are three further varieties, technically known as Arman, Kili, and Kilen, which are also best classified as separate languages, though they are often included in the context of Ewen (Arman) and Nanai (Kili and Kilen), so also in the present volume. Arman may be characterized as an exceptionally archaic variety of Ewen and therefore transitional towards Ewenki, while Kili and Kilen are "mixed" languages, combining Nanai morphology with an Ewenki or (early) Oroch-Udihe type of basic vocabulary. Although extremely interesting in the taxonomic framework, all these varieties are insufficiently documented. Arman and Kili, which were spoken on the Russian side, are recently extinct, while Kilen on the Chinese side is still reported to have a few speakers.

DATA AND SOURCES

Apart from Manchu, the Tungusic languages became first known to international scholarship via the publications of the German naturalists in Russian service in the 18th century. In his comparative lexicon P. S. Pallas (1786–1789) provides lexical data from up to nine varieties of "Tungus", that is, Ewenki and Ewen, as well as Manchu. Thanks to its wide areal distribution Ewenki was the natural object of many of the early attempts of documentation. Material for the first grammatical description of (Khamnigan) Ewenki was collected by the Finnish field linguist M. A. Castrén (1856), as published by Anton Schiefner, who also published several other early collections of data. In the latter half of the 19th century, vocabularies and short grammatical descriptions became available on Oroch and Nanai.

Ever since its beginnings, the field of Tungusic studies has evolved in the context of German and Russian scholarship. The Russian tradition of Tungusic language studies, as summarized by Gorcevskaya (1959), was initially focused on Manchu, but it soon expanded to comprise also several other Tungusic languages, including, in particular, Ewenki, Ewen, Nanai, and Oroch. Some written materials of religious and educational content were also published in Ewenki, Ewen, and Nanai using a Cyrillic-based ortography. The authors of many of the early publications were not professional linguists, with the notable exception of A. O. Ivanovskii, whose data on Solon (1894) retain their value on a dialectal variety that still awaits a more comprehensive documentation. Without going deeper into the individual languages, the posthumously published lexical work of S. M. Shirokogoroff (1944), later re-organized and republished by Doerfer and Knüppel (2004), should also be mentioned since it covers several varieties of both Ewenki and Orochen from a pan-dialectal perspective.

During the Soviet period, the documentation of the individual Tungusic languages was initially stimulated by the need to develop orthographies and literary languages for the

purpose of raising literacy and increasing the political consciousness of the indigenous peoples. While many of the activists of the initial period ended up as schoolteachers, or perished during the purges of the 1930s, some became established scholars. After WW2 and until the 1980s, there was a "Golden Age" dominated by several eminent Soviet Tungusologists, each of whom was typically specialized on one "major" Tungusic language, including G. M. Vasilevich on Ewenki, K. A. Novikova on Ewen, and V. A. Avrorin on Nanai. Avrorin also worked on Oroch and authored a grammar of Written Manchu, published posthumously (2000). There were, of course, also several other scholars, many of whom worked on a number of more "minor" Tungusic varieties, including L. D. Rishes on Arman, T. I. Petrova on Uilta, E. P. Lebedeva (together with V. A. Avrorin) on Oroch, and L. I. Sem on the dialects of Nanai.

Two of the Soviet Tungusologists of the "Golden Age" rose above the others in that they initiated the comparative and diachronic study of Tungusic. V. I. Cincius, who initially worked on Ewen, but later also on Ewenki and Neghidal, authored the first ever "comparative phonetics" of the Tungusic languages (1949). Later, she became the initiator and editor-in-chief of the "comparative dictionary" of the Tungusic languages (1975–1977), which still remains the only work of its kind. O. P. Sunik, whose synchronic work concerned, in particular, Ulcha and Kili, authored comparative monographs on syntax (1947) and morphology (1962, 1982). The works of Cincius and Sunik operate with the entire range of Tungusic languages, including Manchu, as well as with a broad spectre of their dialectal varieties. It may be noted that Russian publications on Tungusic normally apply a Cyrillic-based quasi-phonemic transcription. A more advanced project aiming at creating a unified phonetic transcription for all the Tungusic languages was introduced by Novikova (1961). This transcription, which had both Cyrillic and Roman-based options, was, however, never adopted for general use.

Brief systematic descriptions of all Tungusic languages (with the exception of Manchu), as authored by Russian Tungusologists of the Soviet period, are also present in the collective volume published under the editorship of P. Ya. Skorik in the series "Languages of the Soviet Union" (1968). A similar volume, published in the post-Soviet period, but still involving many Soviet-period authors, was edited by V. M. Alpatov & al. in the series "Languages of the World" (1997). In general, the post-Soviet period meant initially a decline in the Tungusological tradition of Russia. During Soviet times, Tungusic studies had been concentrated in the "Altaic Sector" (*Sektor altaiskikh yazykov*) of the Leningrad Branch of the Institute of Linguistics of the Soviet Academy of Sciences, but due to a series of administrative reforms the "Sector" was closed and its remaining staff dispersed to other entities and institutions. Even so, several of the younger scholars with a background in the "Altaic Sector" have continued their Tungusological activity also later. At the same time, new regional centres of Tungusic studies have developed in places like Novosibirsk and Blagoveshhensk.

The German tradition of Tungusology was in the 20th century continued mainly in the context of "Altaic" studies, though many of the German Altaists have actually been adherents of the Anti-Altaist point of view. After WW2 the field was taken up in Germany by Johannes Benzing, who, on the basis of Soviet primary sources, authored a descriptive grammar of "Lamut" (1955), followed by a "comparative grammar" of the Tungusic languages (1956), which up to the present day remains the principal reference tool on the topic. While arriving on many points at results similar to those of Cincius (1949), Benzing added important theoretical insights and established a base for a Western taxonomical and terminological tradition concerning the grammatical categories of Tungusic.

Other eminent German scholars who started working on Tungusic in the postwar period include the Pro-Altaist K. H. Menges, who contributed a large general survey of the Tungusic languages. combining both synchrony and diachrony, to the Tungusological volume published under the general editorship of Berthold Spuler (1968), and the Anti-Altaist Gerhard Doerfer, who, among other things, published several detail studies (1973, 1975, 1984) concerning the taxonomic position of Kili (Kur-Urmi Nanai). Doerfer also worked broadly on the external connections of Tungusic and initiated a project on "North Asian cultural history", the main result of which was an important collection of papers under the joint editorship of himself and Michael Weiers (1978). The volume also contains a map of the Tungusic languages, compiled and commented on by Robert G. Service (1978).

Another important language of Tungusic studies is Japanese. Modern Tungusology in Japan was founded by Ikegami Jirō, who initially focused on Uilta, a language whose speakers lived partly under Japanese administration on South Sakhalin (*Karafuto*) in 1905–1945, from where they were evacuated to Hokkaido after WW2. Ikegami did, however, work also on Manchu, as well as on general Tungusology. His collected papers, most of which are in Japanese, were published in two volumes (1999, 2001). His student Tsumagari Toshirō continued along similar lines of research, but added a component of field work in China and the Russian Far East. He also contributed important papers on issues connected with the transnational position and endangered status of the Tungusic languages in Russia and China (1996, 1997). Stimulated by opportunities of fresh field work, a new generation of Tungusologists has subsequently grown up in Japan. A channel for the field materials is offered by the series "Publications on Tungus Languages and Cultures" (ツングース言語文化論集), in which some 70 volumes have appeared so far, many of them by Kazama Shinjirō.

Compared with Japan, Tungusology in China has had a late start. Apart from Manchu studies, in which field several major reference works, including dictionaries, have been produced, the Tungusic languages were included in the series "Grammatical Sketches of the Languages of the Minority Nationalities of China" (中國少數民族語言簡志叢書), in which volumes on Solon, Orochen, "Hezhe", and Sibe were published, all in 1986. Importantly, "Hezhe" in this context refers to Kilen, an otherwise little documented language on which the only other general source is the "Hezhen" grammar by Zhang & Zhang & Dai (1989), published exceptionally in English. Otherwise, Chinese scholarship has mainly produced a limited number of relatively unsophisticated collections of primary data from the Tungusic languages spoken in China. Linguistic papers on specific grammatical topics have also been published, especially in the journal "Minority Languages" (民族語文), but they all tend to have the same basic problem in that they are detached from the international framework of Tungusic studies and exhibit an ignorance of the results already achieved elsewhere.

English has not been a major language of Tungusic studies until very recently. An early pioneer in this respect was the Latvian linguist and folklorist Peter Schmidt (Pēteris Šmits), who published lexical data on Neghidal, Ulcha, and Oroch, collected in the early 20th century, in a series of papers (1923ab, 1928) which still retain their value. Apart from occasional papers on specific topics, the Tungusic languages have also been discussed in a number of general introductions to the "Altaic" languages, including those by Nicholas Poppe (1965), Igor de Rachewiltz & Volker Rybatzki (2010), and Václav Blažek & al. (2019). This line of research is being continued today in the framework of "Transeurasian", as displayed broadly in the collective volume edited by Martine

Robbeets and Alexander Savelyev (2020). The volume contains, among other things, grammatical sketches of three sample Tungusic languages, Ewen (Pakendorf & Aralova 2020), Nanai (Oskolskaya 2020), and Sibe (Jang 2020).

During the early stages of the field, the principal task of Tungusic studies was documentation. Today, when many Tungusic languages are kept alive by a rapidly diminishing number of aged semi-speakers, documentation no longer serves linguistic purposes, though it can accompany the efforts aimed at supporting language survival and revitalization where it is still possible. One of the last major documentation projects was initiated by Kim Juwon in Korea, with the aim of recording at least one speaker of each Tungusic (as well as Mongolic and Turkic) language. The results of this extensive enterprise, involving a large number of younger scholars, are being published since 2008 in the "Altaic Language Series" of Seoul National University. Volumes on Ewen, Nanai, and Spoken Manchu (as well on a number of Turkic and Mongolic languages) have already appeared in the series.

Fortunately, most Tungusic languages are well documented, and the extant material offers possibilities for many new types of linguistic work. One of the fields that have not yet been exhausted is areal typology. Among the first to take up this challenge was Li Bing (1996) in his discussion concerning the types and origin of vowel harmony in Tungusic. The first attempt at a typologically oriented description of an entire Tungusic language was the Ewenki grammar by Igor Nedjalkov (1997), followed by the exceptionally voluminous description of Udihe by Irina Nikolaeva & Maria Tolskaya (2001). Although not without problems, these grammars still serve as the main tools on Tungusic for general typologists, who do not have access to the vast literature published in the traditional languages of Tungusology. Specific typological topics have also been taken up by Andrej Malchukov (1996, 2000) and Andreas Hölzl (2015, 2018, 2020). The increasingly international status of Tungusic studies was first publicized in a conference in Bonn (2000) and its proceedings, edited by Carsten Näher & al. (2002–2004). Two other volumes charting the new perspectives of the field include those edited by Malchukov & Whaley (2012) and Hölzl & Payne (2022).

As far as diachrony is concerned, one of the principal issues on which different opinions have been expressed concerns the internal taxonomy of the Tungusic languages. This issue has been touched upon in all general presentations of Tungusic, as well as in many works dealing with the individual languages. Some of the most important specific papers on the issue include Ikegami (1974), Doerfer (1978), Vovin (1993), Whaley & Grenoble & Li (1999), Georg (2004), as well as, most recently, Whaley & Oskolskaya (2020). The languages whose taxonomic position has been viewed as particularly problematic include Manchu, Ewen, and Orochen. Even so, it is possible to arrive at a relatively uncontroversial and generally acceptable classification, as discussed below.

INTERNAL TAXONOMY

All classifications proposed for the Tungusic languages operate with the model of a branching family tree, though in the details, especially with regard to closely related varieties, the model of areal grouping has also been used. It cannot be denied that some boundaries between the individual Tungusic languages are transitional, though others are not. The basic question is, therefore, where the big faultlines within the family go. This, again, depends on how they are defined. Most classifications proposed so far are based on a quantitative analysis of a selected set of features, which can be phonological,

morphological, and/or lexical. In a less mechanistic framework we should, however, take a critical look at the taxonomic parameters, especially in the sense that not all differences and similarities can have the same taxonomic weight. This is because some features are more primary than others, and many features involve secondary parallel developments conditioned by areal contact.

There is a consensus on that the Tungusic language family comprises four clearcut and internally coherent groups (as first explicitly pointed out by Ikegami). The four groups, which in the Japanese tradition of Tungusology are often referred to as groups I, II, III, and IV, will here be termed Ewenic, Orochic, Nanaic, and Jurchenic. The Ewenic group comprises, above all, Ewenki, Ewen, and Solon, but also Arman, Neghidal, and Orochen. The Orochic group, which has also been called "Udegheic", comprises Oroch and Udihe, two languages both of which have historically been identified as "Oroch". The Nanai group comprises Nanai proper, Ulcha, and Uilta, as well as the two "mixed" languages Kili and Kilen. Finally, the Jurchenic group, which has also been called "Manchuric", comprises Manchu and Sibe, as well as their mediaeval ancestor Jurchen. In this context, it is irrelevant whether or not the recorded forms of Jurchen are directly ancestral to the written and spoken forms of Manchu and Sibe, for the Jurchen-Manchu lineage has all the time involved internal dialectal variation, due to which the dialectal base of the normative forms has varied.

Geographically, the four groups form a somewhat unbalanced system. Ewenic covers most of Siberia between the Ob-Yenisei region and the Pacific coast, extending to the Arctic coast in the north, as well as to the northern parts of Manchuria and Mongolia in the south. Jurchenic has likewise a wide distribution, covering historically much of southern and central Manchuria, but also parts of northern and western Manchuria. In addition, the secondary diaspora group of the Jurchenic (Manchu) speaking Sibe was formed in the 18th century in Jungaria, accompanied by a smaller group of Ewenic (Solon) speakers, known as the Ongkor Solon. By contrast, Nanaic is restricted to a rather well-defined territory in the basin of the Lower Amur and its tributaries, including the lower courses of the Sungari and Ussuri, while Orochic occupies the immediately adjacent mountain region of Sikhote-Alin to the east, located between the Lower Amur main basin and the Pacific coast.

There is less consensus on how the four groups are organized with regard to each other in the framework of a binary family tree. Most classifications proposed so far operate with the concept of two principal branches, identified as "Northern Tungusic" and "Southern Tungusic". There is also agreement on that Ewenic belongs to the context of Northern Tungusic, while Jurchenic is Southern Tungusic. For Orochic and Nanaic, with or without Kili and Kilen, however, three different classifications have been proposed. The traditional view is that they form a separate "Amur Tungusic" branch, also termed "Central Tungusic" (Doerfer) or "East Tungusic" (Vovin). This classification is supported by the many structural features which these languages have in common, and which distinguish them from both Ewenic and Jurchenic. Alternatively, Amur Tungusic has been viewed as a separate sub-branch of Southern Tungusic, leaving Ewenic alone in the Northern Tungusic branch (so, for instance, de Rachewiltz & Rybatzki). The third option is that Amur Tungusic is divided between the two principal branches, with Orochic going together with Ewenic into the Northern Tungusic branch and Nanaic together with Jurchenic into the Southern Tungusic branch (Georg).

The division of Amur Tungusic between the Northern and Southern Tungusic main branches, which is the view adopted here, is supported by the distribution of several

primary innovations, especially in the phonology. Thus, for instance, Proto-Tungusic **ü* has a dual representation in the modern Tungusic languages, appearing as (*)*i* in Ewenic and Orochic and as (*)*u* in Nanaic and Jurchenic. Similarly, Proto-Tungusic initial **p* is represented as (*)*p* in Nanaic and Jurchenic vs. (*)*x* in Ewenic and Orochic. Both of these, as well as several other differences involve very early innovations in one or the other branch, or in both, and should therefore be considered as taxonomically diagnostic, whereas, for instance, the morphological features shared by Orochic and Nanaic must be connected with secondary areal interaction. It may be concluded that Amur Tungusic retains its relevance as an areal concept, but not as a primary genetic node in the family tree (Table 1.2).

TABLE 1.2 CLASSIFICATION OF THE TUNGUSIC LANGUAGES

Northern Tungusic		Ewenic	Siberian	Ewen	
				Arman	
				Neghidal	
				Siberian Ewenki	
				Khamnigan Ewenki	Urulyunggui
			Manchurian		Borzya
				Orochen	
				Solon	
			diaspora	Ongkor Solon	
	Amur Tungusic	Orochic		Oroch	
				Udihe	
Southern Tungusic		Nanaic		Nanai	
				Ulcha	
				Uilta	
			mixed	Kili	
				Kilen	
		Jurchenic	mediaeval	Jurchen	
			modern	Manchu	
			diaspora	Sibe	

Within each of the four branches there are important differences between the languages with regard to their degree of innovativeness. This is particularly clear in the Orochic branch, where Udihe may be understood as a highly innovative form of Oroch, with which it is linked via transitional dialects. Similarly, in the Nanaic branch, Nanai is in some respects the most conservative member, while Ulcha and Uilta share a number of secondary innovations absent in Nanai, and Uilta contains an additional set of innovations peculiar to it alone. The Ewenic branch, which has more members, has a somewhat more complex structure, in that it is divided into two major sub-branches, which may be termed Manchurian Ewenic and Siberian Ewenic. Siberian Ewenic comprises, most importantly, Siberian Ewenki, but also Ewen and Neghidal, which may be seen as more innovative members of the sub-branch. Ewen is linked with Siberian Ewenki via Arman, a transitional idiom that is best understood as a conservative variety of Ewen.

Manchurian Ewenic, on the other hand, comprises Solon and Orochen, of which the latter shows transitional features towards Siberian Ewenki.

Occasional attempts at revising the internal structure of Ewenic by assigning special taxonomic properties to either Ewen (Vovin) or Orochen (Whaley & Grenoble & Li) have not yielded generally acceptable results. An important transitional position within Ewenic is, however, formed by Khamnigan Ewenki, which is simply the complex of Ewenic dialects spoken by the Khamnigan, whose other ethnic language is Khamnigan Mongol. It happens that the Khamnigan speak two different varieties of Ewenic, one of which, known as the Borzya (or Man'kovo) variety, is a conservative type of Manchurian Ewenic, while the other, known as the Urulyunggui (or Urul'ga) variety, belongs to the context of Siberian Ewenic. Even so, the two varieties of Khamnigan Ewenki share also a number of common features, which must be connected with their mutual areal interaction.

When it comes to innovativeness, Manchu, and Jurchenic as a whole, is traditionally viewed as the most innovative form of Tungusic. By a mechanic count of distinctive properties (Doerfer) Manchu is certainly the most "different" Tungusic language. The main characteristic of Manchu, as compared with all other Tungusic languages, is its preference for analytic solutions instead of the more synthetic structures present in the other three groups, including also Nanaic. It remains a matter of some dispute as to how "original" the synthetic morphology of Nanaic, Orochic, and Ewenic is, but at least many of the synthetic features present in them and absent in Manchu, such as, for instance, the system of person markers, are clearly primary and have been lost in Jurchenic. This means that the innovativeness of Manchu is mainly due to losses that have simplified the structure of the language in almost all respects, including phonology, phonotactics, morphology, and morphosyntax.

The taxonomic position of the two "mixed" languages, Kili and Kilen, remains also disputable. Earlier classified as Nanai "dialects", they are today generally recognized as separate languages which combine Nanai morphology and morphosyntax with a Northern Tungusic basic vocabulary. Moreover, while the Northern Tungusic lexical component in Kili (as first identified by Sunik) is of a Manchurian Ewenic type, the Kilen lexicon contains some features that would seem to be more compatible with Orochic (as argued by Kazama). In this situation, it remains a matter of opinion whether these languages are to be understood as varieties of relexified Nanai or restructured Northern Tungusic. Altogether, although relatively much research has been done on, especially, Kili, the taxonomy and mutual relationships of Kili and Kilen are a field that still remains to be studied in much more detail.

TYPOLOGICAL PROFILE

All Tungusic languages are unambiguous members of the Transeurasian typological belt traditionally known as "Ural-Altaic", or, in a more restricted framework, as "Altaic". Tungusic is also geographically adjacent to all the other language families belonging to this belt, that is, Uralic (in the west), Turkic (originally in the southwest), Mongolic (in the south), Koreanic (in the southeast), and Japonic (historically on Sakhalin). Bilingualism in either Turkic (Yakut) or Mongolic (Daghur, Khamnigan Mongol, mainstream Mongolian) has been characteristic of certain Tungusic speaking groups for centuries. Therefore, it is not surprising that most of the typological features observed in Tungusic are also present in the other "Ural-Altaic" languages. There are, however, differences in

the details, and some features are areally transitional both within Tungusic and within the "Ural-Altaic" complex.

The following is a list of some of the most prevalent typological features present in Tungusic. The list proceeds from segmental phonology and phonotactics (1–4) to morphology (5–8) and syntax (9–12). As the discussion is focused on structural properties, no actual language data are quoted in this context.

(1) Consonants: All Tungusic languages have a relatively simple consonant system comprising most typically 18 consonants, and distinguishing between four places of articulation (labial, dental, palatal, velar), combined with six to seven manners of articulation (nasals, weak stops, strong stops, fricatives, glides, and liquids). The stops (weak vs. strong) are differentiated according to either voice (voiced vs. unvoiced) or aspiration (unaspirated vs. aspirated), or both. The system of liquids comprises a lateral (*l*) and a vibrant/trill (*r*), but there is a clear areal tendency to reduce the role of the latter in the languages spoken in the Amur region. The system of fricatives comprises typically a dental sibilant (*s*) and a velar continuant (*x*), but some languages have additional segments in the labial and palatal sets. Other additional consonants include a set of distinctive postvelars, present marginally in Manchu, but phonetically also in several other Tungusic languages in the Amur region.

(2) Consonant phonotactics: In initial (#C) and final (C#) positions only single consonants are permitted, and medial clusters can contain only two segments (C_1C_2), with the restrictions concerning their combinations varying from language to language. Geminate consonants (C_1C_1) are generally rare, and tend to be simplified if they arise at morpheme boundaries. A few languages, notably Ewen, Solon, and Oroch, have, however, secondary geminates produced by processes of vowel elision and/or consonant assimilations. Of the individual consonants the vibrant (*r*) does not appear in initial position, although the lateral (*l*) does. Most Tungusic languages have also an initial velar nasal (*ŋ*), a Siberian areal feature absent in most other languages of the "Ural-Altaic" belt. The phonotactic status of the velar continuant (*x*), often realized as a laryngeal spirant, varies depending on its origin: in the individual languages, it can represent an "original" initial consonant of the same type (**x*), an initial strong labial stop (**p*), or an initial or medial dental sibilant (**s*).

(3) Vowels: The system of vowels is more variable and comprises variously from 5 (in Manchu) to 9 (in Solon) distinctive short vowels, organized in 2–4 levels of articulation. Most Tungusic languages, with the exception of Manchu, have also a varying number of long vowels, which may be analysed as vowel sequences (V_1V_1). Diphthongoid sequences of two non-identical vowels (V_1V_2) are present especially in Amur Tungusic. Phonational distinctions (breathy vs. creaky voice), reminiscent of tones, are present in Udihe, but otherwise suprasegmental features (other than vowel length) are not distinctive in Tungusic.

(4) Vowel phonotactics: The distribution of vowel qualities is regulated by a progressive vowel harmony, which in most Tungusic languages operates on the vertical (high-low) axis, with or without an accompanying tongue root opposition (advanced vs. retracted tongue root). In a few marginal idioms (western Ewen), however, vowel harmony operates on the horizontal (palatal-velar) axis, opposing front vowels to back vowels. Most Tungusic languages have also traces of a labial harmony, due to which the non-high unrounded back vowel (*a*) cannot be combined with the corresponding rounded vowel (*o*). Both types of vowel harmony affect both word roots

and suffixes. Originally phonetic phenomena, they have gained a phonological status due to secondary developments, such as the neutralization of the opposition between certain vowel qualities. These developments differ from language to language.

(5) Parts of speech: There is a clear morphological and syntactic contrast between two word classes, which may be identified as nominals and verb(al)s. In this context, nominals comprise the subclasses of nouns, spatials (spatial nouns), adjectives, numerals, and pronouns, all of which have slight morphological idiosyncracies. In addition, there is a small class of invariables with no productive morphology. In terms of root structure, nominals and verb(al)s are similar, with most roots being bisyllabic vowel stems ((C)VC(C)V), but there are also monosyllabic and bisyllabic consonant stems ((C)(VC)VC), though the set of possible final consonants in them is limited to only a few possible segments. In some languages, complications to this system have been created by the secondary loss of final vowels, while in others the system has been simplified leaving only vowel and nasal stems. A small number of roots, functioning as pronouns and other auxiliaries, are monosyllabic with no final consonant ((C)V). Trisyllabic and longer roots are secondary and may be analysed as derivatives.

(6) Morphological strategies: Grammatical and derivational functions are expressed by suffixes, attached to the preceding word root in a fixed order. The coherence of the suffixally marked word is enhanced by vowel harmony, and the morpheme boundary can involve segmental alternations of both the root and the suffix, rendering the sequence less agglutinative and more fusional. The degree of fusional morphology varies, however, being maximal in the northeast (Ewen) and minimal in the southeast (Manchu). There is also a similar areal transition in the general amount of morphology, which increases when moving from Jurchenic via Amur Tungusic to Ewenic. All Tungusic languages have also postclitics, attached to word forms as phonologically dependent but semantically independent elements.

(7) Nominal morphology: The categories of nominal inflectional morphology comprise number, case, and possession. Number marking tends to be obligatory in the Siberian sphere, but less so in the Manchurian sphere. The number of suffixally marked cases varies from 6 (in Manchu and Sibe) to 11 (in Ewen) or 12–13 (in Solon and Siberian Ewenki), though the exact number is always a matter of interpretation. For object marking, most Tungusic languages, except Jurchenic, have two cases (accusative vs. partitive/designative). For locational expressions a three-way contrast between locative (where?), dative/lative (whither?), and ablative (whence?) cases is used, though the distinction between the locative and dative/lative functions is occasionally fuzzy. Most Tungusic languages lack a true adnominal genitive case, though some languages, especially in the Manchurian sphere, have developed a secondary genitive case from what is originally a possessive form (a derivational feature). Personal possession is marked by possessive suffixes of pronominal origin, except in Jurchenic, where only analytic means are used. In the first person plural, a distinction is made between the inclusive and exclusive functions, also in the Jurchenic pronouns, but not in Nanaic (with the exception of Kili and Kilen). A separate set of suffixes expresses reflexive possession. Most Tungusic languages, except Jurchenic, have also a suffixal marker of alienability.

(8) Verbal morphology: The typical verbal paradigm in Tungusic comprises four types of forms: finite indicative forms, finite imperative forms, participles, and converbs. Participles are morphologically nominals, which is why they can take the nominal

type of suffixal marking for number, case and possession. The finite indicative forms take also person marking, except in Jurchenic, but the set of person markers is different from the possessive suffixes. The imperative forms do not generally take suffixal person marking, but they tend to be used in person-specific functions. Most converbs are morphologically invariable, but some can take plural marking of the nominal type. Apart from actual converbs, there is a secondary category of "quasiconverbs", which are morphologically transparent case forms of participles. Many functional categories, including voice, mood, and aspect are expressed by derivational suffixes, located between the verbal root and the inflectional markers.

(9) Phrase structure: The word order in both nominal and verbal phrases is strictly head-final. In sequences of a numeral and a noun, both members can function as head nouns, though more often the numeral functions as a modifier. In the languages with no adnominal genitive, nominal phrases are head-marked, with the nominal or pronominal possessor indicated on the head noun by a possessive suffix. In the languages with both possessive suffixes and an adnominal genitive, double marking can be present, and the genitival attribute precedes an adjectival modifier (GAN). Verbal phrases are, by contrast, dependent-marked, in that the object stands in the accusative (definite object) or partitive/designative (indefinite or designative object) case. Adverbials are also marked by the relevant case suffixes.

(10) Sentence structure: In an independent head clause, all Tungusic languages favour the verb-final basic word order (SOV), though deviations from this are not rare especially in the Siberian sphere, where they may or may not be connected with recent Russian interference. The form of the finite predicate is chosen from the sets of finite indicative and imperative forms, but, importantly, participles can also be used in the finite function with person marking of the possessive type. In practice, there is a tendency observed in all Tungusic languages to replace the finite indicative forms with participles, which, then, function as finite tense forms for the present, past, and (in many languages) future tenses. The finitization of participles, in general typology also known by the somewhat misleading and confusing terms "verbalization" and "insubordination", has in some Tungusic languages led to the synchronic presence of two sets of finitely used forms, in which case the original finite set has tended to develop secondary evidential functions (as in Nanai). In other languages, the original finite set has receded or disappeared altogether (as in Oroch).

(11) Complex sentences: There are no primary conjunctions in Tungusic, but subordinate clauses are linked to the head clause with the help of converbs and quasiconverbs, which express various temporal, conditional, or causal relations. In a complex sentence, the predicate of only the last (main) clause stands in a finite form, which can also be a finitely used participle. Depending on their reference, converbs and quasiconverbs can be divided into the conjunct (same-subject) and disjunct (different-subject) types, though some converbs can be inherently ambivalent. In quasiconverbs this distinction is expressed by the opposition between possessive (disjunct) and reflexive (conjunct) suffixes.

(12) Sentence types: For the equative and existential functions, most Tungusic languages use a fully conjugated copula-existential verb, which, however, can be absent in a temporally and personally unmarked nominal sentence. Negation is primarily expressed by a fully conjugated negation verb, which precedes an invariant connegative form of the semantic main verb. In many Tungusic languages this system has undergone restructuring either by the suffixalization of the conjugated negation

verb to the connegative form of the main verb or by the reduction of the morphology of the negation verb and the creation of a number of invariant negation particles. Existential negation is mainly expressed by a number of language-specific privative nouns ('absent, absence'), which can also be suffixalized. Interrogation in polar questions is expressed by interrogative particles, which are typically (post)cliticized to the finite verb. Some Tungusic languages have a separate corrogative particle, used in non-polar (content) questions containing lexical question words. For other discursive functions, a varying number of discourse particles of the (post)clitical type are used.

GRAMMATICAL FRAMEWORK

The grammatical description of the Tungusic languages is firmly anchored in the research tradition of comparative Altaic studies. This tradition also determines to a large extent the terminology used to describe the grammatical features present in the Tungusic languages. One characteristic of this terminology is that it is generally oriented from form to function, rather than vice versa. This means that a single form, or its cognate representations shared by several languages, is labelled uniformly even though its functions may vary from language to language. This principle also applies within a single language: even if a form has several synchronic functions, it still retains its formal identity and can conveniently always be referred to by the same term. There are, however, exceptions, for in some cases it is motivated to use different terms when referring to the different functions of a single form.

Participles are perhaps the most typical example of multifunctional forms. The forms termed "participles" in Altaic studies are actually polyfunctional nominalized verbs. They can be used in the "prototypical" participial function as adnominal modifiers before a head noun, but they can also be used as independent head nouns indicating either an actor (actor nouns) or an action (action nouns). In the latter function, especially when combined with case suffixes, they serve as "quasiconverbs", that is, as predicates of subordinate clauses. Because of the possibility of finitization they can, however, also have the role of finite predicates. This multiplicity of functions calls for terminological differentiation. When used as nominalized verbs, participles are often assumed to indicate a primarily aspectual opposition, which is why the two basic participial forms are traditionally identified by the terms "imperfective" vs "perfective". However, when used as finite predicates, these forms tend to assume temporal functions and are better referred to by the terms "present" vs. "past". Tense and aspect as functional categories are in these cases, as always, difficult to keep apart. Many Tungusic languages have also a diachronically secondary "futuritive" participle, which, when used in a finite function, may be identified as the "future" tense marker. More rarely do the Tungusic languages offer examples of monofunctional nominalizations, but when such forms are present (as in Manchu and Sibe), they are best referred to as "infinitives", implying nominalized verbs that are only, or primarily, used as independent action nouns.

Another multifunctional verbal form in Tungusic is the so-called "aorist". At a deep diachronic level, this is another participial form, which still has the function of a participle in a few languages (as in Manchu and Solon). More generally, this form has, however, become the marker of a temporally unmarked finite paradigm. In many Tungusic languages, this "aorist" paradigm constitutes the basic "unmarked" set of finite forms, as opposed to the temporally differentiated forms based on finitely used participles. For

this reason, the remaining examples of the participial use of the aorist may be referred to as the "aorist participle". This form has, however, also another use, in that, in a negative clause, it is the form that the main verb takes when combined with the conjugated forms of the negation verb. In this function, the aorist participle is best identified terminologically as the "connegative".

Many terminological problems are connected with the converbs. Unlike the participles and the aorist paradigm, which are relatively uniform in all Tungusic languages, there is much more formal variation in the systems of converbs, and many languages possess language-specific converbs, often recent innovations, that have no direct parallels elsewhere. Nevertheless, it seems reasonable to apply unified names for those converbal forms that are present in two or more Tungusic languages. It may be noted that one of the converbs present in most Tungusic languages is based on the aorist participle and is therefore, on formal grounds, best referred to as the "aorist converb", even though several other names have been used for it in the past. Altogether, the terminology used for the converbs is traditionally less well established and more impressionistic than for most other morphological categories, which is why the creation of a new unified terminology requires an innovative approach. The same is true of many of the modal and aspectual forms which are expressed derivationally, but which, nevertheless, can have cognates in several Tungusic languages.

When it comes to nominal morphology the traditional terminology is better established. For most nominal forms it is reasonable to use maximally simple traditional appellations, although these names do not always cover all of their actual functions. The typical case paradigm of a Northern Tungusic language comprises, for instance, the unmarked nominative, as well as the suffixally marked accusative, partitive, dative, ablative, locative, prolative, directive, and instrumental cases, to which a varying number of more specific secondary cases can be added. The functions behind these labels can, however, be multiple. The dative, for instance, apart from the "prototypical" dative function of indicating the recipient ("indirect object"), has also a wide range of locative and temporal functions, many of which overlap to a varying extent with those of other cases, such as the locative and the directive. In this, as well as in other similar examples, the name of a form should best be understood as a simplified and conventionalized cover term for all the functions that the form has.

REFERENCES AND FURTHER READING

Alpatov, V. M. [В. М. Алпатов] & I. V. Kormushin [И. В. Кормушин] & G. C. Pyurbeev [Г. Ц. Пюрбеев] & O. I. Romanova [О. И. Романова] (eds.) (1997) *Монгольские языки, тунгусо-маньчжурские языки, японский язык, корейский язык* [Mongolic, Tungusic, Japanese, Korean], Языки мира (Languages of the World) [3]: 267–284, Москва [Moscow]: "Индрик".

Avrorin, V. A. [В. А. Аврорин] (2000) *Грамматика маньчжурского письменного языка* [A grammar of Written Manchu], Санкт-Петербург [St. Petersburg]: "Наука".

Benzing, Johannes (1955) *Lamutische Grammatik, mit Bibliographie, Sprachproben und Glossar*, Akademie der Wissenschaften und der Literatur, Veröffentlichungen der Orientalischen Kommission 6, Wiesbaden: Franz Steiner Verlag.

Benzing, Johannes (1956) *Die tungusischen Sprachen: Versuch einer vergleichenden Grammatik*, Akademie der Wissenschaften und der Literatur in Mainz, Abhandlungen der geistes-und sozialwissenschaftlichen Klasse 1955 (11), Wiesbaden: Franz Steiner Verlag.

Blažek, Václav, in collaboration with Michal Schwarz and Ondřej Srba (2019) *Altaic Languages: History of Research, Survey, Classification and a Sketch of Comparative Grammar*, Brno: Masaryk University Press.

Castrén, M. Alexander (1856) *Grundzüge einer tungusischen Sprachlehre nebst kurzem Wörterverzeichniss*, herausgegeben von Anton Schiefner, Nordische Reisen und Forschungen 9, St. Petersburg: Kaiserliche Akademie der Wissenschaften.

Cincius, V. I. [В. И. Цинциус] (1949) *Сравнительная фонетика тунгусо-маньчжурских языков* [Comparative phonetics of the Tungusic languages], Ленинград [Leningrad]: Государственное учебно-педагогическое издательство Министерства просвещения РСФСР.

Cincius, V. I. [В. И. Цинциус] (ed.) (1975–1977) *Сравнительный словарь тунгусо-маньчжурских языков: Материалы к этимологическому словарю* [Comparative diactionary of the Tungusic languages: Materials for an etymological dictionary], vols. 1–2, Ленинград [Leningrad]: "Наука".

Doerfer, Gerhard (1973) 'Das Kur-Urmiische und seine verwandten', *Zentralasiatische Studien* 7: 567–599, Bonn.

Doerfer, Gerhard (1975) 'Ist Kur-Urmiisch ein nanaischer Dialekt?', *Ural-Altaische Jahrbücher* 47: 51–63, Wiesbaden.

Doerfer, Gerhard (1978) 'Classification problems of Tungus', in: Doerfer & Weiers (eds.) (1978), 1–26.

Doerfer, Gerhard (1984) 'Die Körperteilbezeichnungen des Kili (Ein Beitrag zur Frage der dialektologischen Klassifikation)', *Nyelvtudományi Közlemények* 86: 238–246, Budapest.

Doerfer, Gerhard & Michael Knüppel (2004) *Etymologisch-ethnologisches Wörterbuch tungusischer Dialekte* (*vornehmlich der Mandschurei*), Hildesheim: G. Olms.

Doerfer, Gerhard & Michael Weiers (eds.) (1978) *Beiträge zur nordasiatischen Kulturgeschichte*, Tungusica 1, Wiesbaden: Otto Harrassowitz.

Georg, Stefan (2004) 'Unreclassifying Tungusic', in: Naeher & Stary & Weiers (eds.) (2004), 45–57.

Gorcevskaya, V. A. [В. А. Горцевская] (1959) *Очерк истории изучения тунгусо-маньчжурских языков* [A brief history of the study of the Tungusic languages], Серия "История отечественного языкознания" [History of linguistics in Russia] 2, Ленинград [Leningrad]: Государственное учебно-педагогическое издательство Министерства просвещения РСФСР.

Hölzl, Andreas (2015) 'A typology of negation in Tungusic', *Studies in Language* 39 (1): 117–157.

Hölzl, Andreas (2018) 'Constructionalization areas: The case of negation in Manchu', in: Evie Coussé & Peter Andersson & Joel Olofsson (eds.) *Grammaticalization Meets Construction Grammar*, Constructional Approaches to Language 21: 241–276, Amsterdam: John Benjamins.

Hölzl, Andreas (2020) 'The complexification of Tungusic interrogative systems', in: Dankmar Enke & Larry Hyman & Johanna Nichols & Guido Seiler & Thilo Weber (eds.) *Language Change for the Worse*, Chapter 8: i–xxxi, Berlin: Language Science Press.

Hölzl, Andreas & Thomas E. Payne (ed.) (2022) *Tungusic Languages: Past and Present*, Studies in Diversity Linguistics 32, Berlin: Language Science Press.

Ikegami, Jiro (1974) 'Versuch einer Klassifikation der tungusischen Sprachen', in: Georg Hazai & Peter Zieme (eds.), *Sprache, Geschichte und Kultur der Altaischen Völker*,

Protokollband der XII. Tagung der Permanent International Altaistic Conference 1969 in Berlin, Schriften zur Geschichte und Kultur des Alten Orients 5: 271–272, Berlin: Akademie Verlag. Also published in Ikegami (2001), 395–396.

Ikegami, Jirō 池上二良 (1999)『満洲語研究』[*Researches on the Manchu Language*], 東京 [Tokyo]: 汲古書院.

Ikegami, Jirō 池上二良 (2001)『ツングース語研究』[*Researches on the Tungus Languages*], 東京 [Tokyo]: 汲古書院.

Ivanovskii, A. O. [А. О. Ивановскій] (1894) *Mandjurica I: Образцы солонскаго и дахурскаго языковъ* [Manjurica I: Specimens of the Solon and Daghur languages], С.-Петербургъ [St. Petersburg]: Типографія Императорской Академіи наукъ. Reprint: Debter—Deb-ther—Debtelin: Materials for Central Asiatic and Altaic Studies 2, Budapest: Akadémiai Kiadó (1982).

Jang, Taeho (2020) 'Xibe and the Manchuric languages', in: Robbeets & Savelyev (eds.) (2020), 269–287.

Li, Bing (1996) *Tungusic Vowel Harmony: Description and Analysis*, Academisch Proefschrift, Universiteit van Amsterdam, Holland Institute of Generative Linguistics (HIL) Dissertations 18, Amsterdam.

Malchukov, Andrej L. (1996) 'Internal relative clauses in Tungusic languages in a synchronic and a diachronic perspective', *Sprachtypologie und Universalienforschung* 49 (4): 358–362.

Malchukov, Andrej L. (2000) 'Perfect, evidentiality and related categories in Tungusic languages', in: Lars Johanson & Bo Utas (eds.), *Evidentials: Turkic, Iranian and Neighbouring languages*, 441–469, Empirical Approaches to Language Typology, Berlin: De Gruyter Mouton.

Malchukov, Andrej L. & Lindsay J. Whaley (eds.) (2012) *Recent Advances in Tungusic Linguistics*, Turcologica 89, Wiesbaden: Harrassowitz Verlag.

Naeher, Carsten & Giovanni Stary & Michael Weiers (eds.) (2002–2004) *Proceedings of the First International Conference on Manchu-Tungus Studies* (Bonn 2000), vol. 1: *Trends in Manchu and Tungus Studies*, Tunguso-Sibirica 8 (2002), vol. 2. *Trends in Tungusic and Siberian Linguistics*, Tunguso-Sibirica 9 (2004), Wiesbaden: Harrassowitz Verlag.

Nedjalkov, Igor (1997) *Evenki*, Routledge Descriptive Grammars, London and New York: Routledge.

Nikolaeva, Irina & Maria Tolskaya (2001) *A Grammar of Udihe*, Mouton Grammar Library 22, Berlin: Mouton de Gruyter.

Novikova, K. A. [К. А. Новикова] (1961) *Проект единой фонетической транскрипции для тунгусо-маньчжурских языков* [A project on a unified phonetic transcription for the Tungusic languages], Москва [Moscow] & Ленинград [Leninrgad]: Академия Наук СССР, Институт языкознания.

Oskolskaya, Sofia (2020) 'Nanai and the Southern Tungusic languages', in: Robbeets & Savelyev (eds.) (2020), 305–320.

Pakendorf, Brigitte & Natalia Aralova (2020) 'Even and the Northern Tungusic languages', in: Robbeets & Savelyev (eds.) (2020), 288–304.

Pallas, P. S. (1786–1789) *Linguarum totius orbis vocabularia comparativa / Сравнительные словари всѣхъ языковъ и нарѣчій*, vols. 1–2, Petropoli / Въ Санктпетербургѣ: Typis Iohannis Carolii Schnoor.

Poppe, Nicholas (1965) *Introduction to Altaic Linguistics*, Ural-Altaische Bibliothek 14, Wiesbaden: Otto Harrassowitz.

de Rachewiltz, Igor & Volker Rybatzki, with the collaboration of Hung Chin-fu (2010) *Introduction to Altaic Philology: Turkic, Mongolian, Manchu*, Handbuch der Orientalistik VIII.20, Leiden: Brill.

Robbeets, Martine & Alexander Savelyev (eds.), (2020) *The Oxford Guide to the Transeurasian Languages*, Oxford Guides to the World's Languages, Oxford: Oxford University Press.

Schmidt, P[eter] (1923a) 'The language of the Neghidals', *Acta Universitatis Latviensis* 5: 3–38, Riga.

Schmidt, P[eter] (1923b) 'The language of the Olchas', *Acta Universitatis Latviensis* 8: 229–288, Riga.

Schmidt, P[eter] (1928) 'The language of the Oroches', *Acta Universitatis Latviensis* 17: 17–62, Riga.

Shirokogoroff, S. M. (1944) *A Tungus Dictionary*: *Tungus-Russian and Russian Tungus*, Photogravured from the manuscripts, Shinobu Iwamura (ed.), Tokyo.

Service, Robert T. (1978) 'Notes on a map of the Tungusic peoples and their dialects', in: Doerfer & Weiers (eds.) (1978), 52–65.

Skorik, P. Ya. [П. Я. Скорик] (editor-in-chief) & al., *Монгольские, тунгусо-маньчжурские и палеоазиатские языки*, Языки народов СССР [Languages of the Soviet Union] 5, Ленинград [Leningrad]: "Наука".

Spuler, B[erthold] (editor-in-chief) (1968), *Altaistik: Tungusologie*, mit Beiträgen von W[alter] Fuchs & Ivan A. Lopatin & Karl H. Menges & Denis Sinor, Handbuch der Orientalistik I.V.3, Leiden: E. J. Brill.

Sunik, O. P. [О. П. Суник] (1947) *Очерки по синтаксису тунгусо-маньчжурских языков* (Essays on the syntax of the Tungusic languages), Ленинград [Leningrad]: Государственное учебно-педагогическое издательство Министерства просвещения РСФСР.

Sunik, O. P. [О. П. Суник] (1962) *Глагол в тунгусо-маньчжурских языках: Морфологическая структура и система форм глагольного слова* (The verb in the Tungusic languages: The morphological structure and system of forms of the verbal word), Москва [Moscow] & Ленинград [Leningrad]: Издательство Академии Наук СССР.

Sunik, O. P. [О. П. Суник] (1982) *Существительное в тунгусо-маньчжурских языках: В сравнении с другими алтайскими языками* [The noun in the Tungusic languages: In comparison with the other Altaic languages], Ленинград [Leningrad]: "Наука".

Tsumagari, Toshirō 津曲敏郎 (1996)「中国・ロシアのツングース諸語」[The Tungusic languages in China and Russia],『言語研究』*Journal of the Linguistic Society of Japan* 110: 177–191, 京都 [Kyoto].

Tsumagari, Toshirō (1997) 'Linguistic diversity and national borders of Tungusic', in: Hiroshi Shoji & Juha Janhunen (eds.), *Northern Minority Languages: Problems of Survival*, Senri Ethnological Studies 44: 175–186, Osaka: National Museum of Ethnology.

Vovin, Alexander (1993) 'Towards a new classification of Tungusic languages', *Eurasian Studies Yearbook* 65: 99–113, Berlin & London: Eurolingua.

Whaley, Lindsay J. & Lenore A. Grenoble & Fengxiang Li (1999) 'Revisiting Tungusic classification from the bottom up: A comparison of Evenki and Oroqen', *Language* 75 (2): 286–321.

Whaley, Lindsay J. & Sofia Oskolskaya (2020) 'The classification of the Tungusic languages', in: Robbeets & Savelyev (eds.), (2020), 81–91.

Zhang, Yanchang [張彥昌] & Zhang Xi [張晰] & Dai Shuyan [戴淑艳] (1989) *The Hezhen Language* [赫哲语], 長春 [Changchun]: 吉林大學出版社 [Jilin University Press].

CHAPTER 2

EARLY FAR EASTERN SOURCES ON TUNGUSIC

Alexander Vovin

The earliest information pertaining to the Tungusic languages and their speakers is contained in Chinese sources starting with the 1st millennium BZ. The early information is, however, invariably circumstantial, hidden behind obscure ethnonyms and brief mentions concerning the mutual intelligibility or non-intelligibility of the local non-Chinese languages, as spoken to the northeast of China. Substantial data start becoming available only after the rise of the Jurchen of the Jin dynasty (1115–1234), a period from which native materials are also preserved in the Jurchen language and script. From that time on, sources written in the Chinese language offer a gradually growing amount of information on Jurchen-Manchu. This information culminates in the large handbooks, dictionaries, and encyclopedias issued during the Qing dynasty (1644–1911), often in bilingual Chinese-Manchu editions.

The situation is similar for Korea, where, as in China, the primary interest was directed to Jurchen-Manchu, a language for which practical tools were needed for the purposes of translation and interpretation. Most of the extant tools were compiled in the 18th century. An additional object of documentation in Japan was offered by the Tungusic languages spoken on Sakhalin and the Lower Amur, a region to which Japanese traders and explorers travelled regularly since the beginnings of the Matsumae rule of Ezo (1590–1871).

Below is a more detailed discussion of the early Far Eastern sources on Tungusic, as arranged according to their origin (China, Korea, and Japan).

CHINESE SOURCES

Unlike Turkic and Mongolic, Tungusic languages appear in Chinese sources quite late. This reflects the fact that the Chinese provided data in their dynastic histories either on their immediate neighbours (Xiongnu, Xianbei, Ruanruan, Tuyuhun, Turks, Japanese, Koreans, and others), or on more remote regions that were crucial for Chinese foreign policy, such as various polities in the Western Regions (Eastern and Western Turkestan). Apparently, until quite late, Tungusic peoples and languages belonged to neither of these categories, a circumstance that provides some support to the point of view that the homeland of Tungusic was located relatively far in central or northern Manchuria. It is sometimes claimed that the *Yìlóu* (挹婁, Late Han Chinese **ʔɨp-l*[*i*]*o*[B], Early Middle Chinese **ʔjəp-ljuᴮ* ~ **ʔjəp-ləuᴮ*), briefly mentioned in volume 85 of the *Later Han History* (後漢書 *Hòu Hàn Shū*) were Tungusic (Pulleyblank 1983: 446), but the linguistic

DOI: 10.4324/9781315728391-2

evidence for such a conclusion is lacking. The same can be said about the *Mòhé* (靺鞨, Late Han Chinese **mat-gat*, Early Middle Chinese **mwât-ɣât*) tribes, which have also been assumed to have been Tungusic speaking.

There are many scattered Jurchen words in Chinese transcription in various sources starting from the Song period (960–1279), which were carefully studied in an excellent monograph by Sun Bojun (2004). The first coherent linguistic source produced by the Chinese on a Tungusic language is, however, a short vocabulary of Jurchen, *Explanation of the Language of the Jin State* (金國語解 *Jīn guó yǔ jiě*), found as an appendix after volume 135 of the *History of Jin* (金史 *Jīn shǐ*), which was compiled in 1343, that is, during the Yuan period (1271–1367). This source presents Jurchen in Chinese transcription without citing the Jurchen script. Its usefulness is limited, as approximately half of it is taken up by titles and proper nouns. However, in the remaining portion we find some usual Jurchen words, for example, 阿鄰 †*alin* 'mountain'. The interesting aspect of this vocabulary is that it also lists some Jurchen words not found in other sources, for example 兀典 †*ucen* (Early Mandarin **ŭ-tjěn*) 'bright star' vs. usual Jurchen 兀失哈 †*ušiha* (Kane 1989: 136) ~ 斡失哈 †*ošiha* 'star' (Kiyose 1977: 97), or 合喜 †*hahi* (Early Mandarin **xɔ́-xǐ*) 'puppy' vs. usual Jurchen †*indahu* (Kane 1989: 216) ~ †*indahun* 'dog' (Kiyose 1977: 105). While †*ucen* 'bright star' has no explanation and can be a corrupted form, †*hahi* 'puppy' looks like a loan from the Old Korean predecessor of Middle Korean †*kàhí* 'dog'. It should be noted that care must be taken when using the edition of this vocabulary in the *Imperially Commissioned Explanation of Liao, Jin, and Yuan Words* (欽定遼金元三史語解 *Qīndìng Liáo Jīn Yuán sānshǐ yǔ jiě*) from the 46th year of Qianlong (乾隆) reign (1782), as this particular edition introduced numerous mistakes in the process of 'correcting' Jurchen data according to Manchu.

During the following Ming period (1368–1644), two vocabularies of Jurchen were compiled: the *Sino-Jurchen Vocabulary of the Bureau of Translators* (女眞館譯語 *Nǔzhēn guǎn yìyǔ*, NZGYY), and the *Sino-Jurchen Vocabulary of the Bureau of Interpreters* (女眞譯語 *Nǔzhēn yìyǔ*, NZYY). The former was first published by Grube (1896), and later, much more recently, by Kiyose (1977). It was also used by Jin Qizong in his dictionary of Jurchen (1984). The second was published by Kane (1989). The difference between the two vocabularies is that while NZGYY represents the written Jurchen language, the NZYY provides a unique glimpse into the spoken Jurchen of the time. Also, while NZYY does not include any Jurchen script, it is considerably bigger than NZGYY.

With the advent of the Qing dynasty in 1644, there was a surge in the publication of Manchu dictionaries, grammars, textbooks and readers. All these sources are in fact semi-native, and not just Chinese. The main goals of those sources, especially of those published from the 1750s onward, was to help the ethnic Manchu, who were rapidly losing their native language, to retain and improve their proficiency in Manchu, as well as to help ethnic Chinese to learn the language. With few exceptions, most of these materials remain unpublished in modern times and exist only in manuscript or xylograph form.

Among dictionaries the three most important ones are: *Dà Qīng quán shū* (DQQS 大清全書) / **Daicing gurun i yooni bithe** 'A Complete Book of the Great Qing' (1683), the first dictionary of the Manchu language, in two volumes comprising fifteen fascicles; *Qīng wén huì shū* (QWHS, 清文彙書) / **Manju isabuha bithe** 'A Classified Book of the Qing language' (1750), in one volume with twelve fascicles; and **Manju gisun i buleku bithe** (MGBB) 'A Mirror of the Manchu Language', a thematic Manchu-Chinese dictionary, compiled during the second half of the Kangxi (康熙) reign (1662–1722). The last dictionary was subsequently revised and enlarged in 1772 and republished under the

new title *Yù zhì zēng dìng Qīng wén jiàn* (御製增訂清文鑑) / **Han i araha nonggime toktobuha Manju gisun i buleku bithe** 'A Revised and Enlarged Mirror of the Manchu Language Written by the Emperor' (HANTMGBB), which remains the largest of all Manchu-Chinese dictionaries. In the following years of the Qianlong reign it was further enlarged by adding first a Mongolian, then a Tibetan, and finally a Uighur section. This latest pentaglot version, known as *Wǔ tǐ Qīng wén jiàn* (五体清文鑑) / **Sunja gisun i hergen kamciha Manju gisun i buleku bithe** 'A Mirror of the Manchu Language Placed Together in the Letters of Five Languages' (WTQWJ), is probably the best known of all Qing period dictionaries. It was reprinted in a typeset edition by the Nationalities Press (民族出版社 *Mínzú Chūbǎnshè*) first in 1957 and then again in the 1990s in three huge volumes. Later it has also been republished with five indexes by Corff et al. (2013).

Among grammars of Manchu, *Mǎn Hàn zǐ Qīng wén qǐ méng* (滿漢字清文啓蒙) / **Manju Nikan hergen i Cing wen ki meng bithe** 'A Manual of Chinese and Manchu Letters' (MHZQWQM) may be mentioned.

Below is a chronological list of the Chinese sources discussed above (abbreviation, Chinese name in characters, Chinese name in transcription / Manchu name, date of compilation or publication):

HHS	後漢書 *Hòu Hàn Shū*, 5th century AZ
JS	金史 *Jīn shǐ*, 1343
NZGYY	女眞館譯語 *Nǚzhēn guǎn yìyǔ*, Ming period after 1407
NZYY	女眞譯語 *Nǚzhēn yìyǔ*, late Ming period
DQQS	大清全書 *Dà Qīng quán shū* / **Daicing gurun i yooni bithe**, 1683
MGBB	**Manju gisun i buleku bithe**, between 1692 and 1722
MHZQWQM	滿漢字清文啓蒙 *Mǎn Hàn zǐ Qīng wén qǐ méng* / **Manju Nikan hergen i Cing wen ki meng bithe**, 1730.
QWHS	清文彙書 *Qīng wén huì shū* / **Manju isabuha bithe**, 1750
YZDQWJ	御製增訂清文鑑 *Yù zhì zēng dìng Qīng wén jiàn* / **Han i araha nonggime toktobuha Manju gisun i buleku bithe**, 1772
WTQWJ	五体清文鑑 *Wǔ tǐ Qīng wén jiàn* / **Sunja gisun i hergen kamciha Manju gisun i buleku bithe**, between 1772 and 1796

KOREAN SOURCES

Like the Chinese, the Koreans used to be engaged in the study of Jurchen and Manchu, but not of the other Tungusic languages. The history of these studies is described by Hiu Lie (1972), who, however, focuses on Manchu, offering only a very brief account of Jurchen language studies. Fortunately, a much earlier study by Ogura Shinpei contains a succinct account of Jurchen language studies in premodern Korea (1940: 660–670), complemented by a brief but highly informative description of Manchu language studies as well (1940: 603–634). More recently, Seong Baeg-in (1970) has published a description of one of the earliest Manchu inscriptions, *Samjǒndobi* (三田渡碑) 'Samjǒndo inscription' (SJDB), which is now standing in the yard of a residential area in southern Seoul in the open air, fully exposed to the elements of weather (as a result of this "preservation" policy, the inscription is gradually becoming almost illegible). There are also many other contributions by Korean scholars regarding Jurchen and Manchu studies, among which those by Kang Sin-hang (1966) and Song Ki-Joong (1981–1982, 2000, 2001: 156–180) are the most important and representative.

During the Koryŏ dynasty (918–1392) the study of the Jurchen language was chiefly conducted in Tongmungwan (通文館), the official Bureau of Interpreters and Translators of the state. This Bureau prepared and published fourteen different readers in Jurchen (Ogura 1940: 668–669). Unfortunately, not a single one of them is extant today. Should even one of them turn out to have survived, it would become a significant addition to the extremely limited corpus of Jurchen texts. The earliest extant Jurchen inscription in Korea is *Kyŏngwŏnbi* (慶源碑) 'Kyŏngwŏn inscription' (KW), which was originally located in Korea's North Hamgyŏng province (which at the time when it was made was not Korean but Jurchen territory), but which was moved in 1918 (during Japanese colonial rule) to the National Museum of Korea. Apart from this, the earliest Korean text *Yongbi ŏch'ŏnga* (龍飛御天歌) 'Song of Dragons Flying in Heaven' (YŎ), written in the national alphabet *Han'gŭl* (한글) in 1445, preserves several Jurchen words in Korean transcription in the commentary. Unfortunately, practically all of them are either place names or proper nouns.

Manchu studies in Korean fared much better in terms of preservation of sources. Out of multiple readers and dictionaries compiled during the Yi dynasty (1392–1910), a majority is still extant. Dictionaries from this period include: *Tongmun yuhae* (TY, 同文類解), 'Explanation of Identical Words' in two volumes; *Han-Ch'ŏng munkam* (HCM, 漢淸文鑑) 'A Mirror of Chinese and Manchu Words' (HCM), also known as *Han-Han-Ch'ŏng munkam* (HHCM, 韓漢淸文鑑) 'A Mirror of Korean, Chinese and Manchu Words' in fifteen volumes, with a facsimile edition by Min Yŏng-gyu (1956); *Pangŏn cipsŏk* (PC, 方言輯釋) 'A Collection and Explanation of Dialects' in four volumes; and *Samhak yŏg.ŏ* (SY, 三學譯語) 'The Trilingual Translations' (SY) in six volumes. The *Pangŏn cipsŏk* and the *Samhak yŏg.ŏ* also include Mongolian and Japanese. As far as readers are concerned, the largest and most important ones are *Sam.yŏk ch'onghae* (SC, 三譯總解) 'A General Explanation of Trilingual Translations' in ten volumes and *Ch'ŏng.ŏ Nogŏltae Ŏnhae* (CNŎ, 淸語老乞大諺解) 'A Manchu Language *Laojida*' in eight volumes, later published as *Ch'ŏng.ŏ Nogoltae Sinsŏk* (CNSS, 淸語老乞大新釋) 'A New Commentary on the Manchu Language *Laojida*' (CNS) also in eight volumes, as published in a facsimile edition with an index of Korean word forms by Chŏng Kwang (1998, cf. also Tsumagari 1977–1978). Smaller readers of Manchu include *So-a ron* (SAR, 小兒論) 'Small Children's Discourse' and *P'alse-a* (PSA, 八歲兒) 'Eight-year-old Children', each of which has only one volume.

It is difficult to overestimate the importance of the Korean eighteenth-century materials on Manchu, for in them we have foreign transcriptions of Manchu in a language that uses an alphabetic system of writing. Thus, the most important information provided by these transcriptions is on the history of Manchu phonology and phonetics throughout the eighteenth century. It has to be added that while some pioneering research has been done by Korean scholars, much of the research carried out so far has focused on the history of Korean language, rather than of Manchu. Therefore, these materials remain a virtually untouched gold mine for Manchu studies, awaiting a systematic and detailed analysis.

The Korean sources discussed above are listed below in chronological order (abbreviation, Korean name in Chinese characters, Korean name in transcription, date of compilation or publication):

KW 慶源碑 *Kyŏngwŏnbi*, between 1138 and 1153
YŎ 龍飛御天歌 *Yongbi ŏch'ŏnga*, 1445

SJDB 三田渡碑 *Samjŏndobi*, 1639
CNŎ 淸語老乞大諺解 *Ch'ŏng.ŏ Nogŏltae Ŏnhae*, 1703
SAR 小兒論 *So-a ron*, 1720 (?), later edition 1774
TY 同文類解 *Tongmun yuhae*, 1748
SC 三譯總解 *Sam.yŏk ch'onghae*, 1703, second revised edition 1774
CNS 淸語老乞大新釋 *Ch'ŏng.ŏ Nogoltae Sinsŏk*, 1765
HCM 漢淸文鑑 *Han-Ch'ŏng munkam*, 1775 (?)
PSA 八歲兒 *P'alse-a*, 1777
PC 方言輯釋 *Pangŏn cipsŏk*, 1778
SY 三學譯語 *Samhak yŏg.ŏ*, 1789

JAPANESE SOURCES

The Japanese chronicle *Azuma kagami* (AK, 吾妻鏡), 'Azuma Mirror' preserves the earliest known foreign transmission of a Jurchen text in original script, 囪土爻肴昊 †*GURun-ni KADAgun* 'bona fide of the country' (Kiyose 1991: 364–370). Apart from this early curiosity, and unlike the situation in China and Korea, there was never any systematic study of either Jurchen or Manchu in premodern Japan. This, however, is well compensated by a number of unique 18th to 19th-century Japanese sources on the Amur Tungusic languages Uilta (Orok) and Ulcha (Santan). These sources include *Karafuto nishi okuchi risū sho* (カラフト西奥地理数書) 'A Record of a Survey of the Far Western District of Sakhalin' (KNORS); *Hokuchi risū torishirabe sho* (北地理数取調書) 'A Report of a Survey of the Northern Land' (HRTS); *Ezo zōshi* (蝦夷草紙) 'Essays on Ezo' (EZ); *Karafuto-jima zakki* (瓦喇弗吐島雑記) 'Various Notes on Sakhalin' (KZ); *Karafuto-jima chiri Santan Rosia-koku kikigaki* (唐太島地理三鞮魯斎亜國聞書) 'Records of Sakhalin, Santan, and Russia' (KCSRK); *Ezo fudoki* (Seki) (蝦夷風土記—関) 'Ezo Gazeteers (Seki)' (EF-S); *Ezo fudoki* (Niiyama) (蝦夷風土記—新山); 'Ezo Gazeteers (Niiyama)' (EF-N); *Karafuto zakki* (Nakamura) (唐太雑記—中村) 'Various Notes on Sakhalin' (KZ-N); *Hen'yō bunkai zukō* (辺要分界図考) 'A Study on the Geography of Important Outlying Areas' (HBZ); *Ezo go shū* (蝦夷語集) 'A Collection of Ezo Words' (EGS) by Uehara Kumajirō (上原熊次郎), mainly containing data on Ainu; *Shindan Karafuto kaitōki* (新談カラフト廻島記), 'Newly Discussed Records on Going around Sakhalin' (SKK), containing only names of trade items; *Karafuto shima shinsetsu* (哈喇土島新説) 'A New Account of Sakhalin Island' (KSS); *Karafuto nikki* (唐太日記) 'Sakhalin diary' (KN), containing only numerals; and *Santan goi* (山丹語彙) 'A Santan Vocabulary' (SG). There is also a vocabulary collected by Mogami Tokunai (最上徳内) and contained in the manuscript titled "Einige Nachrichten über Krafto und Sandan, ein Auszug aus dem Tagebuch meines alten Freundes Mogami Tok'nai auf seinen Reisen dahien" by Ph. Fr. von Siebold.

These sources were studied in great detail by Ikegami Jirō, the doyen of Tungusic studies in modern Japan (Ikegami 2002abc). Unfortunately, in spite of their relatively large volume, the value of these early materials on Uilta and Ulcha is somewhat reduced by the fact that they are all written in the Japanese *katakana* syllabic alphabet. Due to the syllabic nature of the transcriptions and to the fact that Japanese has much simpler phonotactics than any Tungusic language, most of the transcriptions are imprecise and sometimes simply misleading. For this reason, the Japanese transcriptions should only be used in conjunction with the more recent data available in the Cyrillic or Latin alphabets.

The Japanese sources discussed above are listed below in chronological order (abbreviation, Japanese name in original script, Japanese name in transcription, date of compilation or publication):

AK	吾妻鏡 *Azuma kagami*, 1266
EF-N	蝦夷風土記―新山 *Ezo fudoki* (Niiyama), 1789
EZ	蝦夷草紙 *Ezo zōshi*, 1790
KZ	瓦喇弗吐島雑記 *Karafuto-jima zakki*, 1790
KCSRK	唐太島地理三鞮魯斎亜國聞書 *Karafuto-jima chiri Santan Rosia-koku kikigaki*, 1792
EF-S	蝦夷風土記―関 *Ezo fudoki* (Seki), 1792
KZ-N	唐太雑記 *Karafuto zakki* (Nakamura), 1801
HBZ	辺要分界図考 *Hen'yō bunkai zukō*, 1804
EGS	蝦夷語集 *Ezo go shū*, no date, but probably before 1811
HRTS	北地理数取調書 *Hokuchi risū torishirabe sho*, mid 19th c.
KNORS	カラフト西奥地理数書 *Karafuto nishi okuchi risū sho*, mid 19th c.
KN	唐太日 *Karafuto nikki*, 1854
KSS	哈喇土島新説 *Karafuto shima shinsetsu*, 1854
SKK	新談カラフト廻島記 *Shindan Karafuto kaitōki*, 1854
SG	山丹語彙 *Santan goi*, compiled between 1882 and 1886, but probably containing data collected earlier

REFERENCES AND FURTHER READING

Chŏng, Kwang 鄭光 (ed.) (1998)『清語老乞大新釋』[A new explanation of the Manchu language *Laojida*], Seoul: Taehaksa.

Corff, Oliver & Kyoko Maezono & Wolfgang Lipp & Dorjpalam Dorj & Göröochin Gerelmaa & Aysima Mirsultan & Réka Stüber & Byambajav Töwshintögs & Xieyan Li (eds.) 2013, *Auf kaiserlichen Befehl erstelltes Wörterbuch des Manjurischen in fünf Sprachen—„Fünfsprachenspiegel*", vols. 1–6, Wiesbaden: Harrassowitz.

Grube, Wilhelm (1896), *Die Sprache und Schrift der Jučen*, Leipzig: Kommissions-Verlag von O[tto] Harrassowitz.

Ikegami, Jirō 池上二良 (2002a)「十九世紀なかごろのオロッコ語彙ーサンタン語・ギリャク語を含む」[A Vocabulary of mid-nineteenth century Orok—Including Santan and Ghilyak words], in:『ツングース・満州語資料訳解』[Translation and a commentary to Tungusic and Manchu language materials], 154–260, 札幌 [Sapporo]: 北海道大學出版会.

Ikegami, Jirō 池上二良 (2002b)「サンタンことば集」[A collection of Santan (words], in:『ツングース・満州語資料訳解』[Translation and a Commentary to Tungusic and Manchu Language Materials], 261–321, 札幌 [Sapporo]: 北海道大学出版会.

Ikegami, Jirō 池上二良 (2002c)「サンタンことば拾遺」[An additional collection of Santan words], in:『ツングース・満州語資料訳解』[Translation and a Commentary to Tungusic and Manchu Language Materials], 322–334, 札幌 [Sapporo]: 北海道大学出版会.

Jin, Qizong 金啓孮 (1984)『女真文辞典』[A dictionary of the Jurchen language], 北京 [Peking]: 文物出版社.

Kane, Daniel (1989) *The Sino-Jurchen Vocabulary of the Bureau of Interpreters*, Indiana University Uralic and Altaic Series 153, Bloomington, Indiana: Indiana University, Research Institute for Inner Asian Studies.

Kang, Sin-hang 妻信沆 (1966)「李朝中期以後의 譯者에 對한 考察」[A commentary on the translators from the mid-Yi dynasty period],『成均館大學校論文集』11.
Kiyose, Gisaburō N. (1977) *A Study of the Jurchen Language and Script*, Kyoto: Hōritsubunka-sha.
Kiyose, Gisaburō N. (1991) 'The significance of the new Kitan and Jurchen materials', in:『日本語學とアルタイ語學』[Japanese Linguistics and Altaic Linguistics], 359–377, Tokyo: Meiji shoin.
Lie, Hiu (1972) *Die Mandschu-Sprachkunde in Korea*, Indiana University Uralic and Altaic Series 114, Bloomington: Indiana University Press.
Min, Yŏng-gyu 閔泳珪 (ed.) (1956)『韓漢淸文鑑』[A mirror of Korean, Chinese and Manchu words], vol. 1: Facsimile, vol. 2: Index, Seoul: 延世大學東方學研究所.
Ogura, Shinpei 小倉進平 (1940)『朝鮮語學史』[A history of linguistics in Korea], 東京[Tokyo]: 刀江書院.
Pulleyblank, Edwin G. (1983) 'Chinese and their neighbors in prehistoric and early historic times', in: David N. Keightley (ed.), *The Origins of Chinese Civilization*, 411–466, Berkeley, Los Angeles: University of California Press.
Seong Baeg-in 成百仁 (1970) 三田渡碑滿洲文 [The Manchu inscription of Samjŏn-dobi], *Tong'a munhwa* 東亞文化9: 117–148 [Seoul].
Song, Ki-Joong (1981–1982) 'The study of foreign languages in the Yi dynasty (1392–1910)', I–III, *Journal of Social Sciences and Humanities* 54: 1–45, 55: 1–63, 56: 1–57, Seoul: The Korean Research Center.
Song, Ki-Joong 宋基中 (2000) 「朝鮮時代女眞學/淸學」[The study of Jurchen/Manchu in the Chosŏn dynasty], *Altai Hakpo*『알타이학보』10: 105–119, [Seoul].
Song, Ki-Joong (2001) *The Study of Foreign Languages in the Chosŏn Dynasty* (1392–1910), Korean Studies Dissertation Series No. 1, Seoul: Jimoondang International.
Sun, Bojun 孙伯君 (2004)『金代女真語』[The Jurchen Language of the Jin Period], 瀋陽 [Shenyang]: 遼寧出版社.
Tsumagari, Toshirō 津曲敏郎 (1977–1978)「清語老乞大の研究：満州語研究のための一資料(1), (2)」[A study of *Laojida* in the Manchu language: Materials for the study of Manchu],『札幌商科大学・札幌短期大学論集（人文編）』21: 211–248, 22: 161–169, 札幌 [Sapporo].

CHAPTER 3

EARLY WESTERN SOURCES ON TUNGUSIC

José Andrés Alonso de la Fuente

There is an abundance of Tungusic language material hidden in early western sources dating from the 17th century to the first part of the 19th century. With few exceptions, most of them await to be properly edited according to current philological standards. This task must be approached in tandem with similar initiatives aimed at publishing the even more important unpublished sources from the second part of the 19th century onwards. For illustrative purposes, A. A. Burykin (2006: 52–53) mentions several dozens of unpublished recordings from the early 20th century stored at only a single institution in Russia, the Phonogrammic Archives of the Institute of Literature of the Russian Academy of Sciences.

The present account will focus exclusively on Western sources from the late 17th to the early 19th century. These sources are the result of endeavours initiated and accomplished for the most part by French, Dutch, or German travellers and scholars active in various roles in Imperial Russia. For obvious reasons, the resulting language data usually cover a number of Northern Tungusic lects (mostly Ewenki and Ewen) and/or Manchu. In the following, the focus will be on non-Manchuric sources, for the history of early Manchu studies in Europe is relatively well known (cf. e.g. Söderblom Saarela 2020). The selection of sources included in the list below is inevitably somewhat subjective. For instance, Rüdiger (item #12), whose only contribution, some might argue, was to spread further Strahlenberg's (item #5) mistaken identification of Manchu with Chinese, is included since his publication illustrates in no ambiguous terms the many pitfalls of working with early sources.

To be more precise, the 26-item list below includes the most important works on Tungusic produced in the period of time that goes from 1692, the year in which Nicolaes Witsen's (1641–1717) travelogue was first published, to 1856, when the language data collected by M. A. Castrén (1813–1852) saw finally the day of light. After this period the scene of Tungusic studies was long dominated by Russian (including Soviet) publications. The production of new materials or even the re-edition of previously recorded data outside of Russia comes to a stop during the second half of the 19th century, with only a handful of exceptions, including, notably, the publication in Paris of the collections of Gerhard von Maydell and Alexander Czekanowski by Anton Schiefner (1874, 1877), who had earlier been instrumental in making Castrén's materials available to the scholarly world.

The value of Castrén's work as a *terminus post quem* marking the starting point of the scientific study of the Tungusic languages hardly requires justification. Cincius (1975: xxv) refers to Castrén's data (item #26 in the list below) as the only early

DOI: 10.4324/9781315728391-3

contribution worth being included among the materials of her Tungusic comparative dictionary. There were, of course, plenty of reasons behind this chronological limit: for one thing, the comparative dictionary was primarily based on new field materials collected during the Soviet period, but also, many early western sources were not easily accessible at the time. However, such limitations are not actual any more, which is why it is relevant to take a new look at these sources and to see what they can offer for historical and comparative studies of Tungusic.

TYPES OF SOURCES

The items described below can be classified variously according to their origin, form, or status of publication. With regard to their origin, sources can be either primary or secondary. The former type contains material which, generally speaking, was collected *in situ* either by the author or by a second party, whereas the latter type involves the reproduction of previously published materials, with or without acknowledgement to the source(s), and often in the form of compilation from several different sources, both primary and secondary. Errors of various kinds, including misunderstandings, misreadings, and typos, abound in many sources of this type, and what appear to be conscious modifications, especially with regard to spelling practices, are sometimes introduced without explanation. For example, Klaproth's (item #22) re-use of Messerschmidt's (item #4) data shows numerous discrepancies, as in the first five numerals, which are given by Klaproth as ⟨mûkoon, djuhr, Ilánn, degénn, tóŋa⟩, while Messerschmidt writes ⟨Mukónn, Djuhr, Ilánn, Degénn, Tónga⟩. One of the first tasks of the philologist is to identify and account for these and other such anomalies.

As for the form of the sources, they can be divided into two types: explicit and non-explicit. The explicit type comprises sources that aim at offering language data in a systematic arrangement, for instance, in the form of a wordlist. Such sources are typically the result of organized linguistic fieldwork. Non-explicit sources, by contrast, contain language data, often individual technical terms or random phrases, reflecting casual observations, and scattered throughout the body of a text, or also in the appendices to a text dealing with topics other than language in the first place. Sources of the latter type require the researcher to "recollect" the data and present them in a more transparent (explicit) manner.

With regard to their status of publication, sources are either published or unpublished. The list below contains mainly materials that have been published, either by the original author soon after the date of collection or later by other authors. Publication in this context does not, however, mean the presence of a proper critical edition corresponding to modern philological standards, which is available in only a few cases. Needless to say, there is the possibility that there exist hitherto unknown unpublished sources preserved in manuscript form in various archives, especially in Russia. Even so, it is unlikely that any major early collection of Tungusic data would have remained unknown up to the present day.

Below, the items are grouped on a rough chronological basis into eighteenth-century and nineteenth-century materials. Before Castrén there was little difference in the motives and approaches of the persons doing the documentation, but even so the eighteenth century, dominated by the era of Enlightenment, was characterized by an endeavour to discover the diversity of the world, while the nineteenth century, moving towards the era of

Romanticism, was more focused on the historical connections and origins of phenomena, including languages, a focus that functioned as the ultimate stimulation also for Castrén.

The main goal of the present list is to offer a starting point for future research. It is to be hoped that the ready availability of some of these materials in digitalized form on the Internet today will result in them being properly edited and analysed in the future. Bibliographic information aside, the list below includes information about (i) the type and amount of language material contained in each given source, (ii) any subsequent reproduction of the language material in other publications, indicated by the arrow (→), as well as (iii) any modern philological analysis done on the basis of the source.

EIGHTEENTH CENTURY

1. 1692: Nicolaes Witsen, *Noord en Oost Tartarye, Ofte bondigh ontwerp Van eenige dier landen en volken, zo als voormaels bekent zyn geweest*, vols. 1–2. Amsterdam, 1692; 2nd ed. Amsterdam: François Halma, 1705; reprint Amsterdam: M. Schalekamp, 1785. (References below are provided according to the 2nd edition of 1705.)

- Vol. 1, between p. 6 and p. 7: Manchu fragment.
- Vol. 2, p. 654: Ewenki (Lord's Prayer, absent in the 1st ed. of 1692).
  - → Chamberlain, *Oratio dominica* . . . , p. 14 (item #2).
  - → Leibniz, *Collectanea etymologica* . . . , Pars 2, pp. 374–375 (item #3).
  - → Schultze, *Orientalisch- und Occidentalischer* . . . , p. 108 (item #8).
  - → Adelung, *Mithridates* . . . , vol. 4, pp. 223–225 (item #20).
  - → Marcel, *Oratio dominica* . . . , p. 31 (item #17).
- Vol. 2, p. 678: Ewen numerals.

Remarks: general descriptions claim that this work contains "about four dozen" words in Ewenki and Ewen (see, e.g., Gorcevskaya 1959: 6; Clark 1981–1982: 70–71 mentions only "Ewen"), though there are no wordlists as such. The language data (Ewenki, Ewen) were collected at some point between 1666 and 1677. For the analysis of the Ewenki, Ewen and Manchu data, see Pevnov (2018a, 2018b) and Golvers (2018).

2. 1715: Joanne Chamberlaynio [John Chamberlain], *Oratio dominica in diversas omnium fere gentium linguas versa*. Amstelædami [Amsterdam]: Typis Guilielmi & Davidis Goerei.

Remarks: see item #1.

3. 1717: Godofr. Gvilielmi Leibnitii [Gottfried Wilhelm Leibniz], *Collectanea etymologica illustrationi linguarum, veteris celticae, germanicae, gallicae aliarumque inservientia*, 1–3. Hanoveræ: Sumptibus Nicolai Foersteri.

Remarks: see item #1.

4. 1720–1727: Daniel Gottlieb Messerschmidt, *Forschungsreise durch Sibirien 1720–1727*, Quellen und Studien zur Geschichte Osteuropas [Eduard Winter, Georg Uschmann, Günther Jarosch (eds.)], vols. 1–5. Berlin: Akademie-Verlag, 1962–1977.

- Vol. 2: pp. 72–73 ("Orotong-Tungusicae"), 89–90 ("Tungusische Sprache").
  - → Klaproth, *Asia polyglotta*, pp. 286–288 (item #22).

→ Balbi, *Atlas ethnographique*, p. 37 (as "de la Tungouska-Inférieure") (item #23).

Remarks: According to Messerschmidt's travel diary, one Ewenki helped him to write down the language samples (c. 160 words), which were gathered between June 17 and July 11, 1723.

5. 1730: Philipp Johann von Strahlenberg, *Das Nord- und Östliche Theil von Europa und Asia*. Stockholm: in Verlegung des Autoris.

- *sub* "Harmonia linguarum":

  "Mantischeou" = Chinese

  → Pallas, *Linguarum totius orbis vocabularia* . . . , #170 in Appendix (item #14).
  → Rüdiger, *Grundriß* . . . , pp. 94–95, §195: "Die Sprache der Mantschu" (item #12).

  "Tongusi-konni" = (Spoken) Manchu
  "Tungusi-sabatschi *alias* Lamuti" = Ewen
  "Tongusi-oleni" = Ewenki

  → Schultze, *Orientalisch- und Occidentalischer* . . . , pp. 212–219 (item #8).
  → Hervás, *Aritmetica delle nazioni* . . . , pp. 151–152 [331–332, 336] (item #13).

Remarks: See Krueger (1975). Clark (1977–1978) discusses the (Spoken) Manchu data.

6. 1736–1742: Gerhard Friedrich Müller, *Nachrichten über Völker Sibiriens* (*1736–1742*). Manuscript, only partially published.

Remarks: The Southern Ewenki data from "husshing" dialects (c. 140 words), scattered across Müller's writings, are collected in Helimski & Katz (2003: 237–240).

7. 1742: Jakob Johann Lindenau, *Beschreibung der Peschie Tungusen, oder so genannte*[*n*] *Lamuten zu Ochot*, 1742. Manuscript, only partially published.

Remarks: Lindenau's report contains an Arman-German glossary (550 words). Lindenau refers to the Arman as "Tungusy Udskogo Ostroga"). See Titova (1983: 53–70, Russian translation, and 71–76, Arman-German glossary).

8. 1748: Benjamin Schultze, *Orientalisch- und Occidentalischer Sprachmeister*. Leipzig: zu finden bey Christian Friedrich Gessnern (2nd ed. 1769). p. 204: Ewen ("Lamutische") numerals

→ Hervás, *Aritmetica delle nazioni* . . . , pp. 151–152 [333] (item #13).

Remarks: See items ##1, 5. It is unclear what the provenance of Schultze's Ewen numerals was. They differ significantly from the Ewen numerals found in other early sources (see the comparative table below). It is open to speculation whether Schultze took them from Witsen; he could have "corrected" and modified them to his liking. Witsen confuses '12' and '13', but this is corrected in Schulze. Strahlenberg's '2' is actually '7', while his ⟨Dagalkun⟩ is not even a numeral, cf. ⟨tagalkun⟩