

# A typology of questions in Northeast Asia and beyond

An ecological perspective

Andreas Hölzl

Studies in Diversity Linguistics 20



## Studies in Diversity Linguistics

Editor: Martin Haspelmath

In this series:

1. Handschuh, Corinna. A typology of marked-S languages.
2. Rießler, Michael. Adjective attribution.
3. Klamer, Marian (ed.). The Alor-Pantar languages: History and typology.
4. Berghäll, Liisa. A grammar of Mauwake (Papua New Guinea).
5. Wilbur, Joshua. A grammar of Pite Saami.
6. Dahl, Östen. Grammaticalization in the North: Noun phrase morphosyntax in Scandinavian vernaculars.
7. Schackow, Diana. A grammar of Yakkha.
8. Lijegren, Henrik. A grammar of Palula.
9. Shimelman, Aviva. A grammar of Yauyos Quechua.
10. Rudin, Catherine & Bryan James Gordon (eds.). Advances in the study of Siouan languages and linguistics.
11. Kluge, Angela. A grammar of Papuan Malay.
12. Kieviet, Paulus. A grammar of Rapa Nui.
13. Michaud, Alexis. Tone in Yongning Na: Lexical tones and morphotonology.
14. Enfield, N. J. (ed.). Dependencies in language: On the causal ontology of linguistic systems.
15. Gutman, Ariel. Attributive constructions in North-Eastern Neo-Aramaic.
16. Bisang, Walter & Andrej Malchukov (eds.). Unity and diversity in grammaticalization scenarios.
17. Stenzel, Kristine & Bruna Franchetto (eds.). On this and other worlds: Voices from Amazonia.
18. Paggio, Patrizia and Albert Gatt (eds.). The languages of Malta.
19. Seržant, Ilja A. & Alena Witzlack-Makarevich (eds.). Diachrony of differential argument marking.
20. Hölzl, Andreas. A typology of questions in Northeast Asia and beyond: An ecological perspective.

# A typology of questions in Northeast Asia and beyond

An ecological perspective

Andreas Hölzl



Andreas Hölzl. 2018. *A typology of questions in Northeast Asia and beyond: An ecological perspective* (Studies in Diversity Linguistics 20). Berlin: Language Science Press.

This title can be downloaded at:

<http://langsci-press.org/catalog/book/174>

© 2018, Andreas Hölzl

This book is a revised version of a doctoral dissertation written at the University of Munich (Ludwig-Maximilians-Universität München) that was defended in February 2017.

Published under the Creative Commons Attribution 4.0 Licence (CC BY 4.0): <http://creativecommons.org/licenses/by/4.0/>

ISBN: 978-3-96110-102-3 (Digital)

978-3-96110-103-0 (Hardcover)

ISSN: 2363-5568

DOI:10.5281/zenodo.1344467

Source code available from [www.github.com/langsci/174](http://www.github.com/langsci/174)

Collaborative reading: [paperhive.org/documents/remote?type=langsci&id=174](http://paperhive.org/documents/remote?type=langsci&id=174)

Cover and concept of design: Ulrike Harbort

Typesetting: Andreas Hölzl, Felix Kopecky, Sebastian Nordhoff

Proofreading: Alec Shaw, Alena Witzlack-Makarevich, Amir Ghorbanpour, Benjamin Brosig, Jaime Peña, Jeroen van de Weijer, Linda Lanz, Ludger Paschen, Maksim Fedotov, Martin Haspelmath, Stefan Hartmann, Sune Gregersen

Fonts: Linux Libertine, Libertinus Math, Arimo, DejaVu Sans Mono

Typesetting software: Xe<sub>La</sub>TeX

Language Science Press

Unter den Linden 6

10099 Berlin, Germany

[langsci-press.org](http://langsci-press.org)

Storage and cataloguing done by FU Berlin

Freie Universität  Berlin

Für meine Eltern Margret und Wolfgang



# Contents

Acknowledgments	v
Abbreviations	vii
<b>1 Introduction</b>	<b>1</b>
<b>2 An overview of language families in Northeast Asia</b>	<b>15</b>
2.1 Ainuic . . . . .	19
2.2 Amuric (Nivkh) . . . . .	20
2.3 Chukotko-Kamchatkan . . . . .	20
2.4 Eskaleut (Eskimo-Aleut) . . . . .	21
2.5 Indo-European . . . . .	22
2.6 Japonic (Japanese-Ryūkyūan) . . . . .	25
2.7 Koreanic . . . . .	25
2.8 (Khitano-)Mongolic . . . . .	26
2.9 Trans-Himalayan (Sino-Tibetan) . . . . .	28
2.10 Tungusic . . . . .	30
2.11 Turkic . . . . .	32
2.12 Uralic . . . . .	33
2.13 Yeniseic (Yeniseian) . . . . .	34
2.14 Yukaghiric . . . . .	34
<b>3 Areal typology and Northeast Asia</b>	<b>37</b>
3.1 Theoretical considerations . . . . .	37
3.2 The Eurasian macro-area . . . . .	41
3.3 Mainland Southeast Asia . . . . .	41
3.4 Northeast Asia . . . . .	43
3.5 Subareas in Northeast Asia . . . . .	49
<b>4 The typology of questions</b>	<b>53</b>
4.1 Introduction to the typology of questions . . . . .	53
4.2 Question marking . . . . .	56
4.2.1 Marking strategies . . . . .	61
4.2.2 The scope of question marking . . . . .	70
4.2.3 Interaction of functional domains . . . . .	72
4.2.4 The number of markers . . . . .	75

*Contents*

4.3	Interrogatives . . . . .	76
4.3.1	Semantic scope of interrogatives . . . . .	80
4.3.2	Word class membership of interrogatives . . . . .	84
4.3.3	The diachrony of interrogatives . . . . .	85
4.3.4	Inflectional properties of interrogatives . . . . .	88
4.3.5	Interrogatives and demonstratives . . . . .	89
4.4	Towards an ecological theory of questions . . . . .	90
<b>5</b>	<b>Survey of the grammars of questions in Northeast Asia</b>	<b>103</b>
5.1	Ainuic . . . . .	103
5.1.1	Classification of Ainuic . . . . .	103
5.1.2	Question marking in Ainuic . . . . .	104
5.1.3	Interrogatives in Ainuic . . . . .	109
5.2	Amuric . . . . .	113
5.2.1	Classification of Amuric . . . . .	113
5.2.2	Question marking in Amuric . . . . .	114
5.2.3	Interrogatives in Amuric . . . . .	117
5.3	Chukotko-Kamchatkan . . . . .	120
5.3.1	Classification of Chukotko-Kamchatkan . . . . .	120
5.3.2	Question marking in Chukotko-Kamchatkan . . . . .	121
5.3.3	Interrogatives in Chukotko-Kamchatkan . . . . .	124
5.4	Eskaleut . . . . .	128
5.4.1	Classification of Eskaleut . . . . .	128
5.4.2	Question marking in Eskaleut . . . . .	128
5.4.3	Interrogatives in Eskaleut . . . . .	136
5.5	Indo-European . . . . .	140
5.5.1	Classification of Indo-European . . . . .	140
5.5.2	Question marking in Indo-European . . . . .	140
5.5.3	Interrogatives in Indo-European . . . . .	152
5.6	Japonic . . . . .	165
5.6.1	Classification of Japonic . . . . .	165
5.6.2	Question marking in Japonic . . . . .	167
5.6.3	Interrogatives in Japonic . . . . .	189
5.7	Koreanic . . . . .	200
5.7.1	Classification of Koreanic . . . . .	200
5.7.2	Question marking in Koreanic . . . . .	201
5.7.3	Interrogatives in Koreanic . . . . .	214
5.8	Mongolic . . . . .	217
5.8.1	Classification of Mongolic . . . . .	217
5.8.2	Question marking in Mongolic . . . . .	219
5.8.3	Interrogatives in Mongolic . . . . .	244
5.9	Trans-Himalayan . . . . .	255
5.9.1	Classification of Trans-Himalayan . . . . .	255

5.9.2	Question marking in Trans-Himalayan . . . . .	256
5.9.3	Interrogatives in Trans-Himalayan . . . . .	274
5.10	Tungusic . . . . .	284
5.10.1	Classification of Tungusic . . . . .	284
5.10.2	Question marking in Tungusic . . . . .	286
5.10.3	Interrogatives in Tungusic . . . . .	312
5.11	Turkic . . . . .	331
5.11.1	Classification of Turkic . . . . .	331
5.11.2	Question marking in Turkic . . . . .	333
5.11.3	Interrogatives in Turkic . . . . .	354
5.12	Uralic . . . . .	363
5.12.1	Classification of Uralic . . . . .	363
5.12.2	Question marking in Uralic . . . . .	364
5.12.3	Interrogatives in Uralic . . . . .	371
5.13	Yeniseic . . . . .	377
5.13.1	Classification of Yeniseic . . . . .	377
5.13.2	Question marking in Yeniseic . . . . .	377
5.13.3	Interrogatives in Yeniseic . . . . .	381
5.13.4	Dene-Yeniseian? . . . . .	383
5.14	Yukaghiric . . . . .	386
5.14.1	Classification of Yukaghiric . . . . .	386
5.14.2	Question marking in Yukaghiric . . . . .	386
5.14.3	Interrogatives in Yukaghiric . . . . .	392
<b>6</b>	<b>Interrogative constructions in Northeast Asia: A summary</b>	<b>395</b>
6.1	Question marking . . . . .	395
6.1.1	Marking strategies . . . . .	395
6.1.2	Semantic scope . . . . .	398
6.1.3	Interaction of functional domains . . . . .	399
6.1.4	Borrowing . . . . .	402
6.2	Interrogatives . . . . .	405
6.2.1	Formal properties . . . . .	405
6.2.2	Semantic scope . . . . .	406
6.2.3	Diachrony of interrogatives . . . . .	410
6.2.4	Borrowing . . . . .	412
6.3	The significance of the grammar of questions . . . . .	412
6.4	An atlas of the grammar of questions in Northeast Asia . . . . .	418
<b>7</b>	<b>Conclusion</b>	<b>435</b>
	<b>Appendix A: Data for geographical maps</b>	<b>441</b>
	<b>References</b>	<b>449</b>

*Contents*

<b>Index</b>	<b>501</b>
Name index . . . . .	501
Language index . . . . .	511
Subject index . . . . .	521

# Acknowledgments

This book is a revised version of a doctoral dissertation written at the University of Munich (Ludwig-Maximilians-Universität München) that was defended in February 2017. It was made possible through the support of the Graduate School Language & Literature Munich and especially the German Academic Scholarship Foundation (Studienstiftung des deutschen Volkes).

I would like to express my gratitude to Lindsay J. Whaley for making available to me a conference presentation on Oroqen that was relevant for this study, to Benjamin Brosig for his invaluable comments on the chapter on Mongolic languages, to Andrej Malchukov for some data on the language Even, to András Róna-Tas for discussing some of my thoughts on Alchuka *k*-, to Bernard Comrie, not only for commenting on my typology of questions but also for providing several studies on the languages of Siberia, to Michael Cysouw for allowing the use of his conceptual space of interrogatives, to Erika Sandmann for sending valuable data and explanations on the language Wutun, to Patryk Czerwinski for eliciting some examples from the last speakers of Uilta, to Peter-Arnold Mumm for pointing out shortcomings in the conceptual space of question marking, to Stefan Georg for providing information on Ket and some Mongolic languages, to Marek Stachowski for a brief discussion of details of Turkic interrogatives, and, finally, to Kathleen Rabl for going through my English. I also want to thank my informants of Japanese, Kalmyk, Khakas, Khalkha Mongolian, Korean, Mandarin, Russian, and Xining Mandarin. Needless to say, all the remaining shortcomings are mine. I also want to express my gratitude to Elena Skribnik and especially to Wolfgang Schulze and Hans van Ess for their constant support. Finally, I wish to extend my warmest thanks to Yadi Wu for always being there when I needed help most. Last but not least I would like to thank Martin Haspelmath, Sebastian Nordhoff, the anonymous reviewers, and the proofreaders working for Language Science Press.



# Abbreviations

The glossing of the examples mostly follows the *Leipzig Glossing Rules*. See <https://www.eva.mpg.de/lingua/resources/glossing-rules.php> (Accessed 2016-07-06).

.	old or unimportant morpheme boundary	EX	root expander (Miyaoaka 2012)
~	submorpheme, resonance	EXP	experiential (Sun Chaofen 2006)
#	beginning or end of a sentence	FQ	focus question
ABM	ablative modalis (Miyaoaka 2012)	G	Gyeongsang
APPR	approximative	GQ	grammar of questions
AMH	anatomically modern humans	Got	Gothic
ANE	Ancient North Eurasians	Grk	Greek
APPR	approximative (Shigeno 2010)	HL	highlighter (Stern 2005)
AQ	alternative question	IE	Indo-European
Av	Avestan	INT	interrogative (e.g., <i>who</i> , <i>what</i> )
BG	bound genitive pronoun (Huang 1996)	K	Korean
CAY	Central Alaskan Yupik	KM	<i>kakari musubi</i> (Japanese for <i>focus concord</i> ) (Shinzato 2015)
CDC	Common Dialectal Chinese (Norman 2014)	KP	<i>kakari (musubi)</i> particle (Shinzato 2015)
CIA	Copper Island Aleut	ky	thousand years
cj	conjunct	kya	thousand years ago
CK	Chukotko-Kamchatkan	Lat	Latin
COLL	collective (Kämpfe & Volodin 1995)	MHG	Middle High German
COMM	committal	MK	Middle Korean
CPR	Chinese Pidgin Russian	my	million years
CQ	content question	ms	morphosyntactic separator (Vajda 2004)
CSY	Central Siberian Yupik	MSEA	Mainland Southeast Asia
DJ	disjunct	NAQ	negative alternative question
EC	Early Chinese (Norman 2014)	NE	New English
EOJ	Eastern Old Japanese	NEA	Northeast Asia
		NHG	New High German

## *Abbreviations*

NPQ	negative polar question	PR	Proto-Ryūkyūan
NRF	non-referential (Huang 1996)	PS	Proto-Slavic
OAQ	open alternative question	PT	Proto-Tungusic
OAv	Old Avestan	PQ	polar question
OCS	Old Church Slavonic	p.c.	personal communication
OE	Old English	Q/Q	question
OES	organism-environment system	REL	relevance (Shapiro 2010)
OHG	Old High German	RF	reduced forcefulness (Li & Thompson 1981)
OR	Old Ryūkyūan	SEMF	semi-formal
P	participle	SGS	suggestion (Miyara 2015)
PC	Proto-Chukotian	Skt	Sanskrit
PCK	Proto-Chukotko-Kamchatkan	TA	Tocharian A
PG	Proto-Germanic	TAME	tense, aspect, mood, and evidentiality
PIE	Proto-Indo-European	TB	Tocharian B
PJ	Proto-Japonic	TH	Trans-Himalayan (Sino-Tibetan)
PM	Proto-Mongolic	THM	thematic vowel
PMJ	pre-modern Japanese	TQ	tag question
PN	personal/proper/place name	WOJ	Western Old Japanese
POST	postterminal (Ragagnin 2011: 151)	x	mixed with (Janhunen 2012d)

# 1 Introduction

In recent years the study of linguistic diversity took center stage in linguistic typology (e.g., Evans & Levinson 2009). Nettle (1999: 10) usefully differentiated between three types of linguistic diversity that he called *language diversity* (the number of languages), *phylogenetic diversity* (the number of language families), and *structural diversity* (grammatical differences among languages). This study is concerned with all three kinds of diversity, but places an emphasis on the last. In this it follows Nichols (1992: 2), who postulated that “the main object of description here is not principles constraining possible human languages but principles governing the distribution of structural features among the world’s languages.” Different from a classical and purely synchronic typological study based on a well-balanced global sample of languages, this study openly seeks the areal and genetic bias and investigates the distribution of linguistic and especially of structural diversity in Northeast Asia (NEA). Because “typological distributions are historically grown” (Bickel 2007: 239), this study emphasizes the internal development in individual language families as well as their mutual relations.

The ultimate goal is to understand “**what’s where why?**”, and this makes it clear that the major contributions that typology offers are not confined to Cognitive Science as narrowly understood. The goals of 21st century typology are embedded in a much broader **anthropological perspective**: to help understand how the variants of one key social institution are distributed in the world, and what general principles and what incidental events are the historical causes for these distributions. (Bickel 2007: 248, my boldface)

Bickel (2015) today calls this approach *distributional typology*. Nichols (1992), based on an analogy with biology, employed the term *population typology* instead. Dahl (2001: 1456) prefers yet another name, *areal typology*, defined as “the study of patterns in the areal distribution of typologically relevant features of languages” that “is both descriptive and explanatory” and “has both a synchronic and a diachronic side.” What these approaches have in common is not only their focus on the distribution of diversity, but also the desire to explain its emergence.

The holistic approach taken in this study can be tentatively characterized as an *ecological typology* that is committed to an ecologically plausible understanding of language and human beings (Hölzl 2015b: 186). However, in linguistics *ecology* can be understood in a variety of different ways. So-called *ecolinguistics*, for instance, according to one view “is the study of the impact of language on the life-sustaining relationships among humans, other organisms and the physical environment” and “is normatively orientated towards preserving relationships which sustain life.” (Alexander & Stibbe 2014: 105) In another

## 1 Introduction

sense, the ecological aspect instead refers to the maintenance of languages and ensuing preservation of linguistic diversity (e.g., Mühlhäusler 1992). The approach followed here is less value-driven (Hözl 2015b: 173f.); it concentrates instead on the description and explanation of linguistic diversity. While it shares this focus with the other approaches mentioned above, it emphasizes the importance of ecology for an adequate understanding of language. The fundamental unit of description is the *organism-environment system*, or OES for short (e.g., Turvey 2009; Welsch 2012). According to Järvillehto (1998: 329), the theory of the OES maintains “that in any functional sense organism and environment are inseparable and form only one unitary system. The organism cannot exist without the environment and the environment has descriptive properties only if it is connected to the organism.” This theory has a relatively long history, which is concisely summarized in Järvillehto (2009). For example, Sumner (1922: 233) employed the term *organism-environment complex* instead, but similarly claimed that “the organism and the environment interpenetrate one another through and through.” However, Järvillehto (2009) did not mention a very similar concept called the *life space* advocated by Lewin (1936: 12): “Every scientific psychology must take into account whole situations, *i.e.*, the state of both person and environment.” Language, it will be argued, is an integral component of the human OES. Language is not restricted to the organism (e.g., the brain), but equally has an existence as a self-constructed niche (Odling-Smee & Laland 2009; Sinha 2013), *i.e.* a modification of the environment by an organism such as the web of a spider or the dam of a beaver (Odling-Smee et al. 2013: 5).

Niche construction refers to the modification of both biotic and abiotic components in environments via trophic interactions and the informed (*i.e.*, based on genetic or acquired information) physical “work” of organisms. It includes the metabolic, physiological, and behavioral activities of organisms, as well as their choices.

Human niche construction encompasses a multitude of different examples, ranging from the use of tents such as the Evenki *d’u* (similar to a tipi), over the domestication of reindeer, the construction of railroads, or deforestation, to human-induced climate change. In fact, given the extraordinary impact of humans on the environment, the term *Anthropocene* has been suggested as the contemporary geological epoch (e.g., Rosol & Renn 2017 and references therein). The hypothesis that language is an integral component of the organism-environment system has important consequences for the understanding of linguistic diversity. Of course, linguistic diversity is neither scattered at random, nor is it without limits. Rather, there must be a *reason* for the distribution of linguistic diversity we find today (Bickel 2014; Bickel 2015: 904f.). However, a distinction between synchrony and diachrony is insufficient as a proper explanation. One of the most promising approaches to the *natural causes of language* has recently been put forward by (Enfield 2014: 13ff.), who distinguishes between a total of six *causal frames* in which *linguistic* processes occur.

Each of the six frames – microgenetic, ontogenetic, phylogenetic, enchronic, diachronic, synchronic – is distinct from the others in terms of the kinds of causality

it implies, and thus in its relevance to what we are asking about language and its relation to culture and other aspects of human diversity. One way to think about these distinct frames is that they are different sources of evidence for explaining the things that we want to understand. (Enfield 2014: 13)

These causal frames are related to, but not quite identical with, different time scales, ranging from milliseconds to millions of years (Table 1.1). There is a certain amount of mutual interdependence and influence between these frames, each of which combines properties of both organism and environment to different degrees. Niche construction, for example, may exist at several time scales and can “accumulate over time” (Odling-Smee et al. 2013: 18).

Table 1.1: Examples of causal frames loosely based on Enfield (2014: 13–17) with a focus on language

Frames	Timescales	Examples
phylogenetic	ky-my	biological evolution, climate change, language evolution
diachronic	y-ky	language change, language families, conventionalization
ontogenetic	m-y	individual biography, language acquisition, entrenchment
enchronic	s-m	turn-taking, conversation, question-response sequences
microgenetic	ms-s	physiological processes, action, perception, conception
synchronic	–	language systems, knowledge of a given language

All of these frames are crucial to an explanation of linguistic diversity, although a focus will be on some of them. Originally, linguistic typology was mostly concerned with the *synchronic* dimension, which is a necessary abstraction to consider individual languages as fixed entities that can be described and compared. The *diachronic* frame primarily concerns language change over a period of years or thousands of years. This study in particular investigates what will be called the *grammar of questions* (GQ), i.e. those aspects of any given language that are specialized for asking questions or regularly combine with these.<sup>1</sup> The ability to ask questions as well as the existence of specialized constructions for asking questions seem to be universal. Questions, of course, are part of question-response sequences, which are located in the *enchronic* frame that refers to social interaction. Most theoretical discussions of questions, from a speech act perspective

<sup>1</sup>Cable’s dissertation has the title *The grammar of Q* (Cable 2007). However, the term itself has not been clearly defined and is grounded in generative grammar.

## 1 Introduction

for example, concentrate on this frame (e.g., Levinson 2012a). Exceptions include psychological studies (e.g., Loewenstein 1994) or the so-called *cognitive typology* approach by Schulze (2007), which also include the microgenetic frame. As opposed to the social dimension of the enchronic frame, the *microgenetic* perspective concentrates on the cognitive and physiological processes that take place within the organism-environment system. The emergence of the grammar of questions over *phylogenetic* (human and linguistic evolution) and *ontogenetic* time-spans (individual development, especially of children), as described by Tomasello (2008), will not play an important role in this study.

Apart from the causal frames, it is important to add different *loci of causes*, which can be described metaphorically as different types of ecology that language is embedded in. A recent classification proposed by Steffensen & Fill (2014: 7) distinguishes between four different ecologies:

- (1) Language exists in a **symbolic ecology**: this approach investigates the co-existence of languages or ‘symbol systems’ within a given area.
- (2) Language exists in a **natural ecology**: this approach investigates how language relates to the biological and ecosystemic surroundings (topography, climate, fauna, flora, etc.).
- (3) Language exists in a **sociocultural ecology**: this approach investigates how language relates to the social and cultural forces that shape the conditions of speakers and speech communities.
- (4) Language exists in a **cognitive ecology**: this approach investigates how language is enabled by the dynamics between biological organisms and their environment, focusing on those cognitive capacities that give rise to organisms’ flexible, adaptive behaviour. (my enumeration and boldface)

Of course, a focus on language as such is only an abstraction and the above distinction merely highlights several important perspectives (Steffensen & Fill 2014: 7). Each of the four different ecologies influences all three kinds of linguistic diversity, i.e. language, phylogenetic, and structural diversity.

In many cases the exact influence of the four ecologies is only beginning to be understood (e.g., De Busser 2015), which is why only a handful of examples connected with the grammar of questions can be given here. *Symbolic ecology* refers to the aspect of language contact that has a central position in areal linguistics. It encompasses phenomena such as the borrowing of linguistic items, the creolization of languages, or language shift. For example, many languages of China that share a common Chinese ad- or superstrate have borrowed the question marker *ba* 吧 (see below and §5.9.2.1). *Natural ecology*, too, is an aspect that should not be underestimated (e.g., Axelsen & Manrubia 2014). After all, the distribution of languages even today is determined to a large degree by natural and constructed *affordances*—roughly possibilities of action (Lewin 1936; Gibson 1979)—of our environment such as those of rivers, mountains, roads, bridges, or borders. Climate clearly also influences all three types of linguistic diversity (e.g., Everett et al. 2015; 2016). For example, languages that mark polar questions with intonation exclusively and do not have additional question marking strategies—similar to the total number of languages—strangely cluster around the tropics (Dryer 2013j). In Northeast Asia there are almost no such languages. The *sociocultural ecology* plays an important role in language

spread as well, but also influences the relative prestige and importance of languages. This has a direct influence on language shift and the direction of borrowing of linguistic items in language contact situations. As shown by Trudgill (2011) the *social ecology* can have a strong influence on the complexity of a given language, including aspects of the grammar of questions, such as the interrogative system (see §6.3). Furthermore, the culture and way of life of a speech community may have an impact on the structure of languages. Cysouw & Comrie (2013: 388) argued, for instance, that the languages of hunter-gatherers might have preferences for certain linguistic features such as “relatively many cases of initial interrogatives”, although this could not be confirmed for NEA, which contains few real hunter-gatherer groups and few languages with sentence-initial interrogatives. The last point mentioned, the *cognitive ecology*, especially from a micro-genetic perspective, is an important factor in the structural properties the grammar of questions tends to have cross-linguistically. For example, there is a recurrent structural pattern among many different languages in which a content question is immediately followed by a polar, focus, or alternative question (e.g., *What are you doing, are you crazy?*), which can be explained by aspects of the human conceptual system (see §4.4, §6.3).

In principle, all four perspectives are crucial for a complete investigation of language as well as the grammar of questions. Nevertheless, within this study the focus will lie on the aspect of *language contact* (symbolic ecology). Furthermore, a word of caution is in order. While most scholars would probably agree that there may be fundamental differences among individual symbolic, natural, and sociocultural ecologies, there is often a tacit assumption of the uniformity of human cognition throughout the world. This is what Levinson (2012b: 397) has rightfully called “the original sin of the cognitive sciences—the denial of variation and diversity in human cognition.” In fact, Henrich et al. (2010: 61) have quite convincingly shown that many previous investigations in cognitive science or psychology were strongly biased due to problematic samples of participants that do not accurately represent human diversity. This presents us with a severe problem. For instance, questions, it might be argued, can be seen as a way to verbally resolve curiosity. Problematically, publications on curiosity such as Reio (2011: 453) usually share this tacit assumption of universality:

Curiosity is the desire for new information and sensory experience that motivates exploratory behavior. External stimuli with novel, complex, uncertain, or conflicting properties (i.e., collative stimuli) create internal states of arousal that motivate exploratory behaviors to reduce the state of arousal.

Curiously, there are surprisingly few scientific investigations of curiosity. That is why this study necessarily follows this theory, which is basically a summary of Berlyne (1954; 1960; 1978). But it should be borne in mind that there are personal differences of curiosity in both quantity and quality (e.g., von Stumm et al. 2011).

The bulk of this study is a bottom-up comparison of the grammars of questions in different languages and a tentative explanation of their similarities and differences in terms of some of the causal frames and ecologies sketched out above. As further explained in Chapter 4, the typology of questions proposed in this study will mostly concentrate on

## 1 Introduction

question marking and interrogatives (see also Huang et al. 1999). This is a major difference from previous approaches that are usually based on a distinction between different question types, such as polar and content questions. These two domains—question marking and interrogatives—behave quite differently, for instance as regards the symbolic ecology and diachronic time scale. Interrogatives are known to be generally very conservative (e.g., Diessel 2003). In many instances, an interrogative can even remain stable for thousands of years. For example, English *where* can be directly traced back over a time span of several thousand years to Proto-Indo-European *\*k<sup>w</sup>ór* with the same meaning (Mallory & Adams 2006: 419f.). Proto-Indo-European was probably spoken about 6500 years before present (Anthony & Ringe 2015), which means that the interrogative is *at least* of this age. Diessel (2003: 649) thus correctly concludes that interrogatives (and demonstratives) “are generally so old that their roots are not etymologically analyzable”. Theoretically, similar interrogatives can thus be employed to detect previously unknown old genetic connections between languages. In NEA there are a few possible examples of this sort. The most striking is a personal interrogative ‘who’ that has an uncanny similarity in several families, even if one goes back to the respective proto-languages (e.g., Proto-Mongolic *\*ken*, Proto-Turkic *\*kim* ~ *\*käm*, Proto-Yukaghiric *\*kin* etc.). This will be called the *KIN-interrogative* in this study (see §6.2.1). Furthermore, many languages in NEA have what will be called *K-interrogatives*, that is, they have several interrogatives that share a so-called *resonance* (a submorpheme, see Bickel & Nichols 2007: 209; Mackenzie 2009: 1141) that has the form of a velar or uvular plosive or fricative (e.g., Nanai *xai* ‘what’, *xado* ‘how many’, *xooni* ‘how’). Given its fuzzy boundary and only partly analyzable character, a resonance will be indicated with a tilde (e.g., Nanai *x~*) in order to keep it apart from fully analyzable morpheme boundaries written with a hyphen (e.g., Nanai *xai-wa* ‘what-ACC’). This is similar to well-known submorphemes such as English *gl~*, found in *gleam*, *glimmer*, *glisten*, or *glow*. Despite the fact that the initial consonant cluster is not clearly analyzable, the individual instances nevertheless have a vague similarity in meaning. A resonance usually, but not necessarily, indicates a common origin of different interrogatives within one language. It may be noted, however, that KIN- and K-interrogatives are, first and foremost, typological labels and do not necessarily indicate a common origin of different languages as was assumed by Greenberg (2000: 217–224). They are intended to be analogous to the well-known m-T-pronouns found throughout Eurasia, such as in English *me* and *thee* or Nanai *mi* ‘I’ and *si* ‘you (sg)’ (see Nichols & Peterson 2013). Interrogatives are rarely borrowed, and when they are, this usually indicates an extreme contact situation or perhaps widespread bilingualism. Take Mednyj Aleut, for instance, which may be considered a truly mixed language. It exhibits interrogatives both of Aleut (e.g., *kiin* ‘who’) and of Russian (e.g., *kuda* ‘where’) origin (see §5.4.3). Bickerton (2016 [1981]: 65f.) and Muysken & Smith (1990) argue that creole and pidgin languages may have a preponderance of synchronically analyzable interrogatives such as English *at what time*. Because most languages contain at least some instances of analyzable interrogatives, it will be argued that, in order to identify such instances, the whole *interrogative system* needs to be investigated (Muysken & Smith 1990). In most cases of analyzable interrogatives in NEA the actual interrogative takes first po-

sition (e.g., Manchu *ai-ba-* ‘what-place-’). Generalizing on Bickerton’s (2016 [1981]) and Muysken & Smith’s (1990) assumption, the emergence of several analyzable interrogatives can be said to be an instance of *simplification* in the sense of a “regularization of irregularities”, an “increase in morphological transparency” (Trudgill 2011: 62), and a reduction in the number of actual interrogatives. This is most likely due to a specific type of strong language contact such as massive non-native language acquisition (e.g., McWhorter 2007). In sum, interrogatives may thus indicate different kinds of strong language contact (mixing, simplification) and perhaps very distant genetic relationships. The overall similarity of interrogative systems among related languages can also function as a rough proxy for their time of divergence.

Question marking behaves very differently from interrogatives. Of course, question marking may remain stable over long time spans in some cases, but generally is much less stable and more flexible than the interrogative system and is extremely sensitive to language contact. In NEA alone there are dozens of examples of borrowed question markers. One prominent example is the Chinese marker *ba* 吧 that marks polar questions with an additional moment of supposition (‘isn’t it the case that’). The marker has been borrowed by many languages spoken in China today from diverse language families and in many different regions. Even structural question marking such as verb-first word order as found in Germanic languages has been adopted by some Uralic languages, for example (Miestamo 2011). Question marking thus has the potential to indicate language contact, and this it does quite independently of the intensity of the contact. Even relatively light contact may lead to the adoption of a question marker from other languages. However, question marking cannot suggest distant language families. Without doubt, this difference between the two domains—question marking and interrogatives—is an example of the more general principle “that basic structural features tend to be stable, whereas pragmatically sensitive features such as politeness phenomena and evidentials tend to be unstable.” (Trudgill 2011: 3) But interrogatives and question marking certainly represent the extreme ends of what may be conceptualized as a continuum. More or less, they are in complementary distribution when it comes to genetic inheritance and different types of areal contacts. However, the type of question marking (e.g., initial question marker) appears to be more stable than the actual form of the question marker. For instance, many Tungusic languages have a tendency for sentence-final polar question markers despite the fact that they are etymologically unrelated and attested many thousand kilometers apart, e.g. Sibe =*na*# at the Chinese Kazakh border or Even =*Ku*# in northeastern Siberia. The type of question marking thus seems to take a position between the two extremes. Therefore, the grammar of questions represents an ideal tool for the identification of linguistic convergence, possible middle- or long-range relationships, and instances of unusually extreme language contact. Linguistic diversity, just like archaeological records or the human genome, can thus function as a powerful source for the investigation of human prehistory over time spans of hundreds and thousands of years (e.g., Nichols 1992; Heggarty & Renfrew 2014b). In this study Northeast Asia functions as a testing ground for this tentative methodology (see §6.3).

## 1 Introduction

Northeast Asia (NEA) here is first and foremost defined geographically as the region north of the Yellow River and east of the Yenisei (Figure 1.1). A natural boundary is formed in the north by the Arctic Ocean and in the east by the Pacific. In the northeast, the Bering Strait separates NEA from Alaska. NEA includes all islands along the Pacific Rim up to the Aleutian chain that are all located north of Taiwan, but excludes Taiwan itself, which has stronger ties with Southeast Asia. The islands in the Arctic Ocean are largely uninhabited, which renders them irrelevant for the purposes of this study. The Altai, the Kunlun, the Pamir, the Karakorum, the Tianshan, the Qinling, and the Tibetan Plateau will be taken as natural boundaries to the west, southwest, and south.

Thus defined, NEA is a vast area that covers all of Japan, Mongolia, and the two Koreas as well as all of the Far Eastern Federal district, most of the Siberian Federal district of Russia, and northern China, including Manchuria, Inner Mongolia, Xinjiang, parts of the adjacent provinces, and certain parts of Tibet (Amdo).

Unfortunately, Asia is a clear concept only until one tries to define it properly. It combines cultures and languages as diverse as Israel and the Asiatic Eskimos, it is located on several distinct tectonic plates, the largest of which includes Europe but not India, and there is no meaningful boundary of any sort that would clearly differentiate between Asia and Europe. Thus, in the end one is left with the two possibilities that Sinor (1990) was struggling with when trying to define the cultural area of *Inner Asia*. He was well aware that the term *Inner Eurasia* would have been more adequate, but today the term *Asia* is simply too strongly conventionalized and entrenched. This book similarly makes use of the term *Northeast Asia*, even though *Northeast Eurasia* might have been the better choice. Nevertheless, this makes it compatible with previous approaches with the same name and research on neighboring areas such as *Southeast Asia* (SEA).

Apart from Northern China, Korea, and Japan, NEA is extremely sparsely settled. Even Northeast China (Manchuria) and northern Japan (Hokkaidō) have only been settled in larger numbers within the last 150 years or so (e.g., Janhunen 1996). In contrast with the Western Siberian Lowland and the adjacent regions of European Russia and Eastern Europe, most of NEA may be said to be generally very mountainous or at least to be located at higher altitudes. NEA has important bodies of water, including lakes such as Lake Baikal, which defines something like the center of NEA, and several large rivers that play an important role for the dispersal of languages. In Russia these are, beginning from the west, the Yenisei, the Lena, the Indigirka, and the Kolyma, all of which flow into the Arctic Ocean. Further south, the Amur forms the border between Russia and China before it bends towards the northeast and flows into the Sea of Okhotsk. In China, the Liao flows into the Gulf of Bohai from the north and the Yellow River from the west. There are several smaller rivers such as the Yalu, which forms the border between North Korea and China, or the Anadyr in Chukotka. For the most part, NEA is characterized by a continental climate with cold and often dry winters but warm or hot and more humid summers. However, there are considerable regional differences ranging from a tundra climate in the northern parts of Russia, to a very humid subtropical climate in the south of Japan, to a desert climate in northwestern China as well as parts of Mongolia. The northern parts of NEA are mostly covered by Taiga and, further north, by tundra. As



Figure 1.1: Some natural boundaries of Northeast Asia; adapted from [https://en.wikipedia.org/wiki/Geography\\_of\\_Asia](https://en.wikipedia.org/wiki/Geography_of_Asia), adapted from [http://visibleearth.nasa.gov/view\\_rec.php?vevliid=11656](http://visibleearth.nasa.gov/view_rec.php?vevliid=11656) (Accessed 2016-04-10.)

## 1 Introduction

one moves south, the Taiga changes into mixed forests that give way to the steppes in Inner and Outer Mongolia, the Manchurian and North Chinese Plain, the Ordos Plateau, as well as the deserts Gobi and Taklamakan (e.g., Taaffe 1990; Janhunen 1996; Narangoa & Cribbs 2014).

Parts of NEA have been home to *Homo erectus*, Neanderthals, Denisovans, and possibly to other human (sub)species, the classification of which is still disputed. Despite the possibility that both Neanderthals and Denisovans may have had a language comparable to languages today (e.g., Dediu & Levinson 2013) and the fact that both interbred with modern humans (Sankararaman et al. 2016; Reich 2018 and references therein), there is no direct evidence for the languages these extinct groups may have spoken. For this reason, only the language of anatomically modern humans (AMH) can be investigated here. AMH reached NEA and even the northernmost parts of it at least 45 kya (Pitulko et al. 2016, see also Lbova 2014). However, the earliest records of any language in NEA are from Old Chinese and are only about 3250 years old and thus much younger than Sumerian (about 5000 years old) or Ancient Egyptian (about 4700 years old). If history is defined as that period when written language was present, in large parts of NEA it only started several centuries ago (Bellwood 2013). Linguistic reconstructions of some of the oldest proto-languages located in or close to NEA, such as of Austronesian, Trans-Himalayan (Sino-Tibetan), Uralic, and maybe Dene-Yeniseian, must be several thousand years older than Old Chinese records, but nothing comparable to the time of the first peopling of the area.

The earliest accounts of Northeast Asia such as Nicolaas Witsen's (1705) *Noord en Oost Tartarye* employed the term *Tatary* (or *Tartary*), but were quite inconsistent in their use of it. This name has dropped out of use today and in English there is at present no common designation for what has been defined as NEA above. Only in recent years has there been an increase in the West of publications bearing the name *Northeast Asia* in the title. Interestingly, this is a much more common concept in Japan (*hokutō ajia* 北東アジア), Korea (*dongbuk asia*), Mongolia (*züün xojd azi*), and China (*dōngběi yàzhōu* 东北亚洲), but apparently less so in Russia (*severo-vostochnaja azija*). The origin of the term has recently been concisely summarized by Narangoa & Cribbs (2014: 2):

The term "Northeast Asia" is relatively new. It was introduced into academic discourse in the 1930s by the American historian and political scientist Robert Kerner, who taught at the University of California. Kerner's "Northeast Asia" comprised the Korean Peninsula, the Manchurian Plain, the Mongolian Plateau, and the mountainous regions of Eastern Siberia, stretching from Lake Baikal to the Pacific Ocean.

In her recent book *Early modern China and Northeast Asia*, Rawski (2015) included more or less the same region. My account adds substantial areas to this definition, especially in the north and the west. Nevertheless, my approach is similar to Narangoa & Cribbs's (2014: 2) and Rawski's (2015) in trying to break down traditional conceptions of East Asia and a Sinocentric view. Interestingly, an older definition by Chard (1974: xv), which only came to my attention after the bulk of this study was already written, roughly coincides with my definition above:

The area covered comprises Siberia from the Altai Mountains and Yenisei River valleys eastwards, Mongolia, Manchuria, Korea, and Japan. This area has a certain coherence. Geographically, if we except western Siberia with its close affinities to European Russia, it represents the steppe, forest, and tundra zones of northern Asia, lying beyond the loess farmland of traditional China.

The only difference concerns the exclusion of Xinjiang and other parts of northern China. Xinjiang happens to be included in NEA in this study because of its relatively old ties to central China due to Chinese expansions and trade along the Silk Roads, the presence of a great many northwestern Mandarin speakers today, and some linguistic connections to Amdo and Mongolia. Xinjiang is also included in Nichols's (1992: 25f.) concept of *Northern Asia*, which coincides with my definition, except that it includes those areas between the Yenisei and the Ural Mountains. In his recent book *The peoples of Northeast Asia through time*, Zgusta (2015: 21ff.) is not very clear about his definition of Northeast Asia, but he puts an emphasis on what he calls *Pacific Northeast Asia*, which only includes northern Japan, Sakhalin, eastern Manchuria, Kamchatka, and Chukotka. Here this quite useful term will be adopted to additionally include all of Japan, Korea, and the area around the Gulf of Bohai, i.e. all of insular and peninsular NEA adjacent to the Pacific.

The brief review above is not exhaustive but sufficiently illustrates a wide variety of overlapping designations and definitions of NEA. One of the few authors who draw a more differentiated picture is Janhunen (2010: 284):

In the **widest** sense, Northeast Asia as a geographical and ethnohistorical region can be defined as the entire northeastern part of the Eurasian continent, delimited by the Yenisei in the west and the Yellow River in the south. In the northeast, the region extends, in principle, to the Bering Strait. In a somewhat **narrower** framework, Northeast Asia may be defined as comprising the territory between the Amur and Yellow River basins, including the Korean Peninsula and the Japanese Islands in the Pacific coastal zone, but excluding the northeasternmost limits of what is today the Russian Far East. (my boldface)

This broad definition has clearly been influenced by Chard's point of view (Janhunen 1996: 7). The narrow definition, on the other hand, is more or less identical with the perspective taken by Narangoa & Cribbs (2014) or Rawski (2015) seen above and may be more appropriately termed *Greater Manchuria* instead of Northeast Asia (Janhunen 1996: 6). Needless to say, this study is based on a wide definition of NEA.

The addition of the part *and beyond* to the title of this book has two meanings. First, some languages such as the Turkic languages Chuvash and Turkish that are located outside of, but have ties to, or in these cases even originate in, NEA, will be included as well. This problem of establishing a meaningful western boundary of, in their terms, *northern East Asia* has also been observed by Heggarty & Renfrew (2014a: 873):

Turkish serves also to stress just how far the typological unity of this language area stretches beyond any geographical definition of *East Asia*. For in linguistic terms

## 1 Introduction

– whether in family affiliations, typology or prehistory – northern Asia allows of no meaningful division into eastern or western parts. This language area covers its entirety, westwards to the Urals and, as Turkish (or Finnish) attest, in parts beyond. Its origin and core, however, do lie firmly within our scope here.

Second, despite its focus on one area, this study is still intended to be applicable to other languages. Especially Chapter 4 is a more classical approach to typology that seeks to understand what grammars of questions are cross-linguistically attested and possible (cf. Hölzl 2016b). Therefore, it makes extensive use of data from languages outside of NEA.

The survey of languages in Northeast Asia is intended to be as exhaustive as possible. As Voegelin & Voegelin (1964: 2) put it: “In linguistic ecology, one begins not with a particular language but with a particular area, not with selective attention to a few languages, but with **comprehensive attention to all the languages in the area.**” (my boldface) However, some individual languages are underrepresented because of a lack of data. The accuracy and amount of details of descriptions for languages and families varies considerably with my personal experience and the available literature. This book largely relies on previously published material, but several speakers and experts of individual languages were consulted as well. German examples are based on my knowledge as a native speaker. Given my educational background, literature in Chinese, English, and German form the linguistic core on which this book is based. There are a few French publications on NEA languages that were included as well. Russian and especially Japanese literature was consulted where possible, but not with equal intensity. Therefore, the southern part of NEA is somewhat overrepresented in this study. Finnish, Hungarian, Korean, and Mongolian publications were necessarily excluded. Other languages play no significant role for the study of the languages of Northeast Asia. Unfortunately, most grammatical descriptions are insufficient and only those in English and Japanese usually reach an international standard with adequate analyses of examples and glossing. For a typological study, Chinese descriptions that have a rudimentary glossing with characters but usually lack a clear analysis, are usually more useful than German or Russian publications that, with some exceptions, usually lack glosses or analyses completely. As a consequence, many of the examples found in this study have been painstakingly analyzed by myself as far as possible, by and large following the *Leipzig Glossing Rules*.<sup>2</sup> Remaining uncertainties are signaled with a question mark. For most of the languages in NEA only rather brief accounts are available. These are often limited to mentioning a handful of unexplained interrogatives with very rough translations and, with some luck, unanalyzed examples of polar and content questions. The length of the descriptions of the languages within this study also varies due to extreme differences in the complexity of the grammar of questions. It is not always easy to distinguish between simplicity and a lack of information. But there certainly are extremely complex systems such as in the Yupik languages that require several pages and tables just to give a rough outline. Some of the most complex systems can be found in *Omoti*c languages (Afroasiatic) spoken in Ethiopia (see Amha 2012; Köhler 2013; 2016, and references therein). In comparison

---

<sup>2</sup>See <https://www.eva.mpg.de/lingua/resources/glossing-rules.php> (Accessed 2016-07-06.)

(i.e., relative complexity), most languages of NEA have much simpler and typologically more common grammars of questions (e.g., Miestamo 2008). Given the large number of languages included in this study, the description of individual languages is necessarily somewhat superficial and experts will certainly have a lot more to say about each of them. For several reasons, §5.10 on *Tungusic* is somewhat more extensive than those on other language families. First, my personal knowledge of Tungusic is better than for many other languages in this study. Second, there are extremely good descriptions of questions in some Tungusic languages such as Evenki and Udihe. Third, because of their vast distribution over almost all of NEA, Tungusic could potentially be crucial for this study (see Chapter 3). This study also includes several varieties that were described only from the 1980s onward by Chinese scholars but seem to have mostly gone unnoticed outside of China. Tungusic languages will also sometimes be considered in other chapters to illustrate certain points.

There have been several earlier studies on questions in the languages of NEA. There are many good descriptions of questions in individual languages such as Zhang Dingjing (1991) on Kazakh, M. Hayashi (2010a) on Japanese or Yoon (2010) on Korean, to name but a few examples. There are far fewer studies of questions in more than one language, but still no exhaustive list can be given here. Audova (1997) briefly investigates question marking types in the northern part of NEA, but lacks a clear analysis and confuses interrogative verbs (a subtype of interrogatives) with question marking. Nevertheless, she makes some useful observations on possible areal connections. Luo Tianhua's (2013) dissertation is an investigation of questions in the languages of China and thus covers the southern half of NEA. Unfortunately, the overview of most languages is superficial and not always reliable. For instance, only two and a half pages are devoted to all the Tungusic languages spoken in China (Luo Tianhua 2013: 133–135). Several names of individual languages are erroneous and Korean is wrongly classified as a Tungusic language. Nevertheless, there are useful insights about questions in Mandarin and some other languages. More problematic is Greenberg's (2000: 217–234) investigation of interrogatives in so-called *Eurasiatic* languages, which compares look-alike elements in a more or less random sample of languages and claims to have proven a genetic connection among them. A high-quality description of polar question marking in Uralic languages, on the other hand, some of which are spoken in NEA, is given by Miestamo (2011), which is also the most up-to-date description of polar question marking types. Yet another very good typology of questions in Austronesian languages of Taiwan, mostly excluded from this study, can be found in Huang et al. (1999).

In sum, at its core this study is an investigation of the distribution of structural diversity in the grammar of questions in the limited geographical region of Northeast Asia and beyond. The restriction to one category is necessary for reasons of space and clarity, and the process of zooming in on one region allows a higher resolution and historical accuracy than is usually the case in linguistic typology. Some of the questions addressed by this study are: “What does it mean to question?” (Sanitt 2011: 561) Are questions indeed universal, and if yes, why? What about questions is variable? How can this variation be classified? What are possible motivations behind this variation? What patterns do the

## 1 Introduction

languages of Northeast Asia show with respect to this classification? What roles do geography, genetic inheritance, and language contact play in explaining these patterns? Is there convergent evidence from other disciplines such as genetics? And finally, does the concept of Northeast Asia make sense from the point of view of areal linguistics?

This book is organized into seven chapters, including this Introduction. Chapters 2 and 3 briefly present the languages of NEA from a genetic and an areal perspective, respectively. Chapter 4 introduces a somewhat new typology of questions that is illustrated with languages from around the world. The longest chapter (Chapter 5) gives an extensive overview of the grammars of questions in the fourteen language families of NEA. Readers only interested in the typological aspects are advised to skip over this lengthy chapter and consult Chapter 6 instead, which gives an overview of the findings of the previous chapter, illustrated with several geographical maps inspired by the *World Atlas of Language Structures* (Dryer & Haspelmath 2013). Chapter 7 presents some conclusions, sketches possible avenues for further research, and briefly summarizes the tentative idea of an *ecological typology*. Following the extensive list of References, the Appendix lists the data that were used for the comparative maps of §6.4. At the end of the book there are Name, Language, and Subject Indexes.

## 2 An overview of language families in Northeast Asia

The validity of all fourteen language families of NEA has been proven by means of the classical comparative method. Hammarström et al. (2016) list about 430 different language families worldwide. Of these, Niger-Congo (called “Atlantic-Congo”, 1430 languages) and Austronesian (1274 languages) are, in terms of individual languages, the two largest ones. Indo-European (583 languages) and Trans-Himalayan (475 languages) follow in places three and four. All other families found in NEA are considerably smaller, with several dozen languages at most. As regards the size of the individual languages, i.e. the number of speakers, there are similarly pronounced differences. By counting native speakers only, Mandarin is the largest language worldwide with about one billion speakers. English has less than half the number of native speakers, but including second language learners, it must clearly be considered the largest language in the world, with perhaps up to twice as many speakers as Mandarin. Russian (ca. 150 million, Cribber 2002), Japanese (ca. 130 million, Hasegawa 2015), Korean (ca. 75 million, Song 2005), Ukrainian (ca. 36 million, Young 2006), Uzbek (ca. 20 million, Johanson 2006b), Kazakh (ca. 10 Mio, Muhamedowa 2016), Uyghur (ca. 10 million, Tuohuti Litifu 2012), Mongolian (ca. 5 million, Janhunen 2003e), and Amdo Tibetan (ca. 1.3 million, Ebihara 2011: 42), have more than one million speakers. Of the rest, only Shuri, Yakut, Oirat, Tuvan, and Buryat, and perhaps Santa, have between 200,000 and one million speakers. Most of the remaining languages have well below fifty thousand speakers. But note that several languages, including Mandarin, English, Russian, Ukrainian, Uzbek, and Kazakh, are represented in NEA only by a fraction of the total number of speakers.

The names *Paleo-Siberian* or *Paleo-Asiatic* (*paleoaziatiskije jazyki* in Russian) are sometimes still used as labels for several language families (e.g., Tsumagari et al. 2007), especially Amuric, Chukotko-Kamchatkan, Yeniseic, and Yukaghiric, sometimes expanded to include Ainuic. But this label should be avoided whenever possible, as it does not refer to any valid genetic, areal, or typological grouping.

Ainu, Korean, Nivkh, and sometimes even Japanese, are considered to be linguistic *isolates* that are not related to any other known language. However, the difference between a language isolate and a language family is a matter of degree rather than kind. Historically, an isolate *necessarily* is part of a larger stock that has already disappeared, or the relationship to other languages is too remote to be detectable. A case in point is the language Ket. It is known to be part of the Yeniseic language family, but is its sole survivor. Recent years have seen the rise of the so-called *Dene-Yeniseian hypothesis*, which claims a genetic connection between Yeniseic and the Na-Dene languages in North

America. Without the historical attestation of now extinct varieties of Yeniseic, neither the Yeniseic language family nor its connection to Na-Dene would be known today, and Ket would simply count as a linguistic isolate. Japanese is certainly not an isolate, but together with the Ryūkyūan languages forms the Japonic or Japanese-Ryūkyūan language family. In addition, Ainu, Korean, Nivkh, and Japanese all have a certain amount of internal diversity that is usually described as dialectal variation. Given the absence of any clear definition of what characterizes a language as opposed to a dialect, a clear distinction between an isolate and a language family cannot be drawn. In order to make the description analogous to the other language families, the designation of the language families of Ainu, Korean, and Nivkh will be Ainuic, Koreanic, and Amuric (Janhunen 1996), respectively.

A special group of Northeast Asian languages is formed by several *pidgins*, *creoles*, and *mixed languages*. Their classification is open to debate and depends on the theory of genetic relatedness one adopts (Operstein 2015: 1–3). The pidgins, both of which are extinct by now, were called Govorka (Taimyr Pidgin Russian, Russian x Nganasan), and Chinese Pidgin Russian x Chinese). Both are strongly based on Russian, which is why they will be treated together with the other Indo-European languages (§§2.5, 5.5). Mixed languages include Copper Island Aleut (Aleut x Russian) and Eynu (Uyghur x Persian). For practical purposes these will be treated together with Eskaleut (§§2.4, 5.4) and Turkic (§§2.11, 5.11), respectively. An Ainu-Itelmen hybrid will not be included as it is extinct and has not been recorded to a sufficient degree (Fortescue 2003: 81). Yilan Creole, the only language of Taiwan included in this study, is basically Japanese (§§2.6, 5.6), but has been strongly influenced by Austronesian languages. The status of several varieties in the Amdo Sprachbund, especially Gangou, Hezhou, Tangwang, and Wutun (all Sinitic x Turkic x Mongolic x Tibetic), remains somewhat unclear. But there are some indications that they are creolized varieties of Sinitic and thus will all be treated together with Trans-Himalayan (Sino-Tibetan, §§2.9, 5.9). Several languages, including Alchuka, Bala, Kili, Kilen, and Ussuri Nanai, are to different degrees a mixture of several Tungusic languages and therefore treated in §2.10 and §5.10 on Tungusic.

The Indo-European languages Latin, Sanskrit, and Prakrit as well as the Semitic languages Arabic, Aramaic, and Hebrew, all of which were at some point used as literary languages in parts of NEA, will be excluded. The two Indo-European languages Dutch and Portuguese had only a short-lived and, at least for the purposes of this study, unimportant presence in the maritime southeast of NEA. Today, globalization brings many different languages from all around the world into NEA, especially the larger cities in the south. But apart from English, these languages will be neglected, too. NEA may have been home to languages and whole language families that have disappeared without leaving any records. Some of them may be accessible through the study of loanwords. A case in point is the hypothetical language of the Rouran empire (柔然, 330-555 CE) around Mongolia, for which Vovin (2004) has collected a small amount of material. He concludes that it is probably not related to any surrounding language known to us today. Unfortunately, almost nothing is known about its grammatical structure, let alone its grammar of questions. Another language or family of languages that apparently has

disappeared without trace (Fortescue 2013) was presumably spoken by the recently discovered *Paleo-Eskimos*.

Paleo-Eskimos likely represent a single migration pulse into the Americas from Siberia, separate from the ones giving rise to the Inuit and other Native Americans, including Athabaskan speakers. Paleo-Eskimos, despite showing cultural differences across time and space, constituted a single population displaying genetic continuity for more than 4000 years. On the contrary, the Thule people, ancestors of contemporary Inuit, represent a population replacement of the Paleo-Eskimos that occurred less than 700 years ago. (Raghavan, DeGiorgio, et al. 2014: 1020)

This is by no means the only prehistoric population that is attested in NEA, but the recency of their spread would in principle make them accessible with the standard tools of historical linguistics. Recently, genetic studies came to the conclusion that not only populations in Chukotka, but also Kets, Nganasans, Selkups, Yukaghirs (Flegontov et al. 2016), and speakers of Eskaleut and Na-Dene languages (Reich 2018: 175, 183) are genetically related to the Paleo-Eskimos. It would be tempting to connect this evidence with the Dene-Yeniseian hypothesis (Vajda 2010), but thus far we cannot bring together the linguistic and genetic data as there are too many possible variables. It has by now been demonstrated that not only the Paleo-Eskimos, but in fact all native American populations can be traced back to Asia. In other words, all extant and innumerable extinct indigenous American languages necessarily have their origin in NEA in prehistoric times. The so-called *Beringian Standstill Model* assumes that a population had lived relatively isolated in Beringia, now mostly covered by water, before entering the Americas when the glaciers were on their retreat and the sea levels started to rise (e.g., Moreno-Mayar et al. 2018). Llamas et al. (2016: 1), based on genetic evidence, recently argued “that a small population entered the Americas via a coastal route around 16.0 kya, following previous isolation in eastern Beringia for ~2.4 to 9 thousand years after separation from eastern Siberian populations.” (corrected) In other words, the predecessors of most native American languages—possibly excluding speakers of Na-Dene, hypothetical Paleo-Eskimo, and Eskaleut, all of which spread over North America much later—were still around in Beringia, arguably a part of NEA back then, as recently as 16,000 years ago. It is plausible to assume that this Beringian area harbored a certain amount of linguistic and genetic diversity. For example, there is evidence for a population that today only left some genetic traces in Amazonia and is more closely related to Australasians (see Reich 2018: 176-181 and references therein). This time depth of up to 24,000 years of separation of Siberian and these early native American populations lies well beyond the perhaps 10,000 or so years that are, given ideal circumstances, accessible by means of the comparative method. This means that, from a purely linguistic point of view, generally only a fraction of prehistory, namely the Holocene (from ca. 9,500 BCE, Bellwood 2013: 5f.), is actually accessible. Even so, the age of most language families in NEA is considerably lower and does not even approach that age. The data in Table 2.1 are only approximations and different authors give different estimates. The data quoted were chosen because their point of view seems to be by and large the most accurate according to my current understanding.

## 2 An overview of language families in Northeast Asia

Table 2.1: Approximate rounded age and homeland of the 14 language families; arrows indicate the possible location of the pre-proto languages

Family	Estim. age	Location	Source
Trans-Himalayan	?9000-8000	?eastern Himalayas	Blench & Post 2014
Indo-European	6500	north of Black Sea	Anthony & Ringe 2015
Eskaleut	5000	Chukotka → Southern Alaska	Fortescue 2013
Uralic	5000	western Siberia	Janhunen 2009
Chukotko-Kamchatkan	4000	Kamchatkan isthmus	Fortescue 2005
Japonic	2200	southwest Korea → Honshū	Janhunen 2010; Sean & Toshikazu 2011
Turkic	2200	south Siberia, Mongolia	Yunusbayev et al. 2015
Tungusic	2000	middle Amur	Pevnov 2012
Yukaghiric	?2000	Baikal → middle Lena, Indigirka	Häkkinen 2012; Maslova 2003a
Yeniseic	1500-2200	northern China → south Siberia	Vajda 2004; Vovin et al. 2016
Ainuic	1300	Honshū → Hokkaidō	Sean & Hasegawa 2013
Mongolic	800	northeastern Mongolia	Janhunen 2003d
Amuric	500-1500	Upper Amur	Fortescue 2011; Janhunen 2010
Koreanic	500-1000	southeast Korea	Janhunen 2010

§2.1 to §2.14 will briefly introduce all 14 language families of NEA in alphabetical order. Details of the internal classification of the language families, as well as their grammars of questions, will be described in Chapter 5.

## 2.1 Ainuic

Bugaeva (2012: 463) estimates that there are about 100,000 ethnic Ainu, of whom only a handful still speaks the language. Historically, there are three major groups of dialects, the Sakhalin dialects, the Kuril Islands dialects, and the Hokkaidō dialects (e.g., Bugaeva 2012: 461). Proto-Ainuic has roughly been dated “to the last centuries of the first millennium A.D.” (Vovin 1993: 155). The spread of the three branches probably started in northern Hokkaidō (Sean & Hasegawa 2013) and covered a vast area reaching Sakhalin in the Northwest and the Kuril Islands and maybe even the tip of southern Kamchatka in the Northeast. Today, most Ainu have shifted to Japanese and the last speakers are only found on Hokkaidō. Most of the Sakhalin Ainu moved to Japan after the Second World War and the Kuril Island Ainu were relocated as early as 1884. Both groups of dialects are extinct today. Genetic research has revealed that the Ainu are the result of an admixture from the continental Okhotsk people (perhaps connected to the Nivkh) into the Satsumon population, which itself goes back to the Jōmon population (Takehiro et al. 2007). It is known through the study of place names in the Tōhoku region of Honshū that speakers of Ainu or a language closely related to Ainu once must have lived there as well. According to Bentley (2008b: 33), Chinese recordings of Yamatai toponyms, presumably located in southern Japan, are predominantly Japanese, but may also contain several Ainuic elements. The most likely scenario that also takes recent genetic studies into consideration (Jinam et al. 2012), is that the Ainu, because of the arrival of the Japonic-speaking Yayoi people in Honshū, migrated from Honshū to Hokkaidō, where they mixed with people from the Amuric speaking Okhotsk population, but preserved their language and subsequently spread to the surrounding regions (Sean & Hasegawa 2013: 5). Up to this point in time, no genetic connections of Ainuic with other languages or language families have been proven. The best but still not absolutely convincing attempt to clarify the prehistory of the Ainu language has perhaps been made by Vovin (1993: 175), who could “definitely say that Proto-Ainu is unrelated to any of the neighbouring languages.” He proposed a possible connection with Austroasiatic but this is not generally accepted. Hirofumi & Oxenham (2013: 219) summarized research on the origin of the Jōmon population and concluded “that it ultimately derived from the modern human colonizers of Late Pleistocene Southeast Asia and Australia, who subsequently mixed with later migrants from the northern part of East Asia during the early Jōmon period (c. 12-7 kya) or before”. This would be in accordance with Vovin’s claim of a southern origin, but given the great time depth of the Jōmon culture of 12 ky and the extremely shallow time depth of Ainuic, no further hypothesis can be drawn on possible linguistic connections. For the time being, Ainuic has to be recognized as a stock on its own, but with possible connections to Mainland Southeast Asia and beyond.

The contact languages of Ainuic were Japonic in the South, and Amuric in the North (e.g., Vovin 2016). There is also strong contact to Russian as well as the Tungusic language Uilta on Sakhalin and, on the southern tip of Kamchatka, to Itelmen. Ainu used to be a lingua franca in southern Sakhalin during the 19th century, and was even used by the Japanese (Yamada 2010: 65).

## 2.2 Amuric (Nivkh)

The designation Amuric has been introduced by Janhunen (1996) to refer to the language family to which Nivkh, previously called Gilyak, belongs. The internal diversity appears to be similar to that of Ainuic, with some dialects being mutually unintelligible (Gruzdeva 1998: 7). No relation with other languages has been proven, although Fortescue (2011) recently argued for the possibility of a remote relationship with Chukotko-Kamchatkan languages, which has yet to be verified. There are at most a few hundred speakers left out of a population of a few thousand. Amuric has often been linked with the Okhotsk culture (5th to 13th century AD), which reached as far as Hokkaidō and the Kuril Islands (Fortescue 2011) and had a strong impact on the Ainu (see also Vovin 2016). Based on evidence from the cultural lexicon, Janhunen (2010: 294) assumes an origin of Amuric further to the south in central Manchuria. However, this contradicts both the assumption that Tungusic was spoken along the middle Amur (§2.10) and the hypothesis that the Okhotsk culture was Amuric-speaking. Today, Nivkh is spoken along the mouth of the Amur and in some villages on Sakhalin and perhaps by a few speakers who were resettled in Hokkaidō after the last world war (Fortescue 2016: 1ff.).

Nivkh had intense contacts with several Tungusic languages (e.g., Gusev 2015b) both at the lower Amur (e.g., Negidal, Ulcha), and on Sakhalin (Uilta, Sakhalin Evenki), where there was also contact with Ainuic and, for a short period, with Japanese (see also Yamada 2010). In addition, there is some evidence for old contacts between Amuric and Ainuic (see Vovin 2016). The most important contact language today is Russian, and most Nivkh have switched to speaking Russian.

## 2.3 Chukotko-Kamchatkan

The status of Chukotko-Kamchatkan (or Luoravetlan) as a language family is not recognized by some authors, notably Georg & Volodin (1999). But Fortescue (2003; 2005; 2011) has quite convincingly shown that it has a firm basis. The language family falls into two major branches, Itelmen (Kamchadal) on the one hand and a more diverse branch including Chuckchi, Alutor, Koryak (Nymylan), and Kerek, on the other hand. All scholars agree that Chuckchi, Alutor, Koryak, and Kerek are related, and the controversy surrounds the question of whether Itelmen belongs to the same language family or not. Concerning the origin of Chukotko-Kamchatkan (CK), Fortescue (2005: 3) assumes the following scenario.

The linguistic “centre of gravity”—suggesting the original CK “homeland”—lies around the Kamchatkan isthmus [...], an area presumably reached from the west along the coast of the Okhotsk Sea long before the introduction of the reindeer-herding from further west within the last thousand years or so [...]. The time at which proto-CK may have been spoken in this general area by hunters of wild caribou has been estimated as somewhere around four thousand years [...]; this coincides with the beginnings of the Neolithic cultures of Tarya on Kamchatka and (a little later) Ust-Belaya on Chukotka.

In agreement with an original location further to the west and perhaps to the south, Fortescue (2011) has recently argued for an old genetic relation of Chukotko-Kamchatkan with Amuric, which seems possible but remains to be verified. A recent genomic study has shown that the Chukchi derive about 40% of their genome from a back-migration of a native American population to Asia (Reich 2018: 184). If the same is true for all Chukotko-Kamchatkan-speaking populations, this opens up the possibility that Pre-Proto-Chukotko-Kamchatkan, or a contact language thereof, can be traced to North America.

Two historically attested dialects of Itelmen as well as Kerek have already disappeared, and all the remaining languages except for Chukchi, which has about 10,000 speakers, are highly endangered. Concerning the lifestyle of the speakers of Chukotko-Kamchatkan, Anderson (2006a: 416) mentions an interesting split.

Along the coasts, Chukchi people live as sea mammal hunters, like the local Yup'ik populations, but they live as reindeer herders in the interior. Approximately three-quarters of the Chukchi live as reindeer herders. Northern Kamchatkan groups mainly practice reindeer-oriented economies and fishing and sea mammal hunting along the coasts. The Itelmen live primarily as subsistence fishers.

The herding of reindeer must be a relatively recent innovation brought to the North-east of NEA by other people from the west, but may have been the driving factor in a secondary expansion of Chukchi.

Chukotko-Kamchatkan languages had contact mostly with Even, parts of Yupik, Yukaghiric, Russian and, less importantly, English. Itelmen seems to have had contact with Ainuic as well.

## 2.4 Eskaleut (Eskimo-Aleut)

Eskaleut languages are for the most part not spoken in NEA, but in Alaska, Canada, and Greenland (e.g., Berge 2006). The primary split is between Eskimo and Aleut, the former having an additional division between Yupik, Inuit, and perhaps Sirenikski (e.g., Fortescue et al. 2010: x). In this study only those Eskaleut languages spoken in or in the vicinity of NEA will be included. These are Sirenik(ski), which is extinct, and Naukan(ski) Yupik on the mainland, Central Siberian Yupik on St. Lawrence Island, and Aleut as well as Mednyj Aleut on the Aleut Islands. The languages have all reached their present location from Alaska, where the homeland of Eskaleut was probably located. Very early, at least several thousand years ago, the Aleut started migrating along the Aleut islands towards Asia.

It can only be surmised that the movement that separated Aleut from Eskimo occurred soon after the first arrival of the Eskimo-Aleut family in Alaska over Bering Strait, at least four thousand years ago and some two thousand years before the Inuit-Yupik split. The linguistic evidence suggests at least two major phases here—an ongoing spread westwards as far as the outermost Near Islands (reached some 2,500 years ago), overlaid in more recent times (only a few hundred years ago) by

a wave bearing specifically Eastern Aleut influence from the Alaskan peninsula. (Fortescue 2013: 344f.)

The best known and most important expansion of Eskimo was about a thousand years ago to northern Canada and Greenland. But there were migrations on the Asian side as well, which are more important for the present study (Fortescue 2004).

On the Asian side of Bering Strait, at approximately the same time as the Thule migration eastward from North Alaska, a westward expansion of Punuk culture whaling people probably speaking Central Siberian Yupik was initiated. This eventually reached as far as the Kamchatkan isthmus in the 15th century, as linguistic evidence suggests, although the Eskimo presence must have been short-lived or absorbed by maritime Koryaks and—especially—Kereks (Fortescue 2013: 344)

It is, of course, generally accepted that Pre-Proto-Eskaleut had been located on the Asian side before crossing over to Alaska, but according to Berge (2010: 558) and Fortescue (2013) this must have been at least 4000 years ago. The possible existence of a few Eskimo loanwords in Tungusic languages cannot change that basic fact (cf. Vovin 2015).

Fortescue (2013: 344) hypothesizes that Sireniki may be “a pocket of archaic Eskimo much influenced by Chukchi.” Aleut probably had contact with unknown languages in Alaska and perhaps the Aleut Islands. Both Aleut and Yupik as spoken in Asia had strong contact with Russian and, less importantly, with English.

## 2.5 Indo-European

Indo-European is the most widespread and the largest language family worldwide in terms of speakers. About one third of the global population speaks an Indo-European language. Proto-Indo-European was presumably located on the Pontic-Caspian steppe, perhaps about 4500 BCE (Anthony & Ringe 2015), although there are competing but in my eyes much less likely hypotheses, for example of a location in Anatolia south of the Black Sea (e.g., Heggarty 2013). There is convergent evidence from the human genome, archaeology, and linguistics for the location on the Pontic-Caspian steppe (e.g., Anthony 2007; Allentoft et al. 2015; Anthony & Ringe 2015; Haak et al. 2015; Jones et al. 2015). According to one prominent view, the subsequent spread and the divergence of Indo-European branches can be summarized as follows:

**Archaic Proto-Indo-European** (partly preserved in Anatolian) probably was spoken before 4000 BCE; **early Proto-Indo-European** (partly preserved in Tocharian) was spoken between 4000 and 3500 BCE; and **late Proto-Indo-European** (the source of Italic and Celtic with the wagon/wheel vocabulary) was spoken about 3500-3000 BCE. Pre-Germanic split away from the western edge of late Proto-Indo-European dialects about 3300 BCE, and Pre-Greek split away about 2500 BCE, probably from a different set of dialects. Pre-Baltic split away from Pre-Slavic and other northwestern dialects about 2500 BCE. Pre-Indo-Iranian developed from a

northeastern set of dialects between 2500 and 2200 BCE. (Anthony 2007: 82, my boldface)

Indo-European has a dozen major branches, four of which have, or formerly had, representatives in Northeast Asia as defined here: Tocharian, Iranian (part of Indo-Iranian), (East) Slavic, and (West) Germanic. Historically speaking, Indo-European languages entered Northeast Asia at roughly three different times.

Pre-Tocharian, which may have branched off from Indo-European about 5300 years ago (before all other branches except Anatolian), probably reached the Altai mountains shortly afterwards and is associated with the Afanasievo culture (ca. 3300-2500 BCE) (Mallory 2010: 51; Anthony & Ringe 2015: 208). The Afanasievo culture showed a southward expansion, which would explain why Tocharian is only attested further south in the Tarim basin in two different forms known as Tocharian A (East) and B (West) (e.g., Winter 1998). There are indications of the existence of a third language (Tocharian C), which is attested exclusively in loanwords (Mallory 2010: 48f.). Tocharian has been extinct for at least a thousand years.

Tocharian A, found in documents near Turfan and Qarashähär, and Tocharian B, found mainly around Kucha in the west but also in the same territory as Tocharian A. The documents, dating from the 6th to the 8th centuries CE, suggest that Tocharian A was by that time probably a dead liturgical language, while Tocharian B was still very much in use. In addition to Tocharian, administrative texts have been discovered in Prakrit, an Indian language from the territory of Krorän [lólán 楼兰]; these documents contain many proper names and items of vocabulary that would appear to be borrowed from a form of Tocharian (sometimes known as Tocharian C) spoken by the native population. The Kroränian documents date to ca. 300 CE and provide our earliest evidence for the use of Tocharian. For our purposes here, it is also very important to note that the earliest evidence for the mummified remains of “westerners” in the Tarim Basin is found in cemeteries at Xiaohe [小河] (Small River) and Qäwrighul [gümügöu 古墓沟], both of which are located in the same region as Tocharian C. (Mallory 2010: 48f., my square brackets)

There are alternative names for Tocharian A, such as *Agnean* after the Sanskrit name Agni (*yānqí* 焉耆) for the city of Karashahr, and for Tocharian B, such as *Kuchean* after the city of Kucha (*qiūzī* 龟兹 and variants) (e.g., Fortson 2010: 400; Geng Shimin 2012).

Tocharian was in contact with several Iranian languages that entered the Northeast Asian scene after Tocharian, but were probably present in the Tarim basin as early as 1300 BCE (Mallory 2010: 50). Iranian together with Indic and maybe Nuristani as an independent subbranch, forms the Indo-Iranian branch of Indo-European (Fortson 2010: 202f.). Iranian language history is usually divided into an Old Iranian (until the 4th or 3rd century BCE), a Middle Iranian (until the 8th or 9th century CE), and a Modern Iranian period (e.g., Schmitt 2000: 3). Iranian languages only had a wide distribution in NEA during the Middle Iranian period. The two languages Khotanese (*hétíán sàiyǔ* 和田塞语, in the South of the Tarim basin, ca. 5th to 10th century CE, Emmerick 2009:

377ff.) and Tumshuquese (*túmùshūkè sàiyǔ* 图木舒克塞语, in the North, 7th to 8th century CE), closely related and usually collectively called Saka (*sàiyǔ* 塞), were more restricted in their distribution than Sogdian (Emmerick 2009; Geng Shimin 2011). Sogdian (*sùtèyǔ* 粟特语, ca. 4th to 11th centuries CE, Yoshida 2009: 279ff.) was originally spoken in present-day Uzbekistan and Tajikistan, but “the Sogdians played an active role as international traders along the Silk Road between China and the West, with the result that the Sogdian language became a kind of lingua franca in the region between Sogdiana and China” (Yoshida 2009: 279). Regarding modern Iranian, only the Pamir languages Sarikoli (*sàlikù'er* 萨里库尔) and Wakhi (*wāhān* 瓦罕), treated as dialects of one language called *tǎjíkèyǔ* 塔吉克语 (Gao Erqiang 1985: 101) but not to be confused with the Tajik language, as well as the mixed Persian-Uyghur language Eynu, are spoken in NEA. However, the discussion will also briefly mention Yagnobi, which is located in Tajikistan but represents the only modern language that is closely related with Sogdian.

The last period of Indo-European influx brought Eastern Slavic as well as Germanic languages into NEA. Together with the Baltic languages, Slavic forms the Balto-Slavic branch of Indo-European. Only the East Slavic languages Russian and Ukrainian expanded into NEA. Russian is not only the dominant language of the Russian Federation, but has also had some influence on several languages outside of Russia, such as Mongolian or Uyghur. Many speakers of languages in the Russian Federation are bilingual in Russian or are even shifting to Russian as their primary language. Ukrainian only plays a marginal role, but nevertheless can be found scattered across the Russian-speaking area. Slavic originates in Eastern Europe, perhaps northwest of the Black Sea (Fortson 2010: 420f.) and the Russian expansion beyond the Urals only started in the 16th century. By 1625 the Russians reached the Yenisei, and by the end of the 17th century they had conquered most of Siberia, excluding only Outer Manchuria, Chukotka, and southern Kamchatka (Forsyth 1992: 102). This means that Russian played no role in NEA until about 400 years ago. There is a mixed Russian-Ukrainian language called Surzhyk, of which some speakers are most likely also found in NEA, but which must be neglected for lack of sufficient information (Bilaniuk 2004).

Only West Germanic languages are marginally represented in NEA by scattered minorities of German (especially Altai Low German) speakers living in southern Siberia as well as a certain amount of influence from American English as spoken in Alaska and the Aleut Islands. Yiddish is included here mostly because of the existence of a Jewish Autonomous Oblast in Russia close to Khabarovsk, where a handful of Yiddish speakers can be found and where it has an official status. Yiddish is a descendant of primarily southeastern Middle High German that was extensively influenced by Slavic, Hebrew, and Aramaic (Jacobs et al. 1994). Altai Low German (or Plautdiitsch) “is the descendant of the Low German (Low Prussian and Pommeranian) dialects once spoken in the Danzig area.” (Nieuweboer 1999: 13) There is only limited information on questions in Altai Low German, but Standard German, a liturgical language for Siberian speakers of German dialects, can give some rough indications about how the blanks may be filled in. There was an English jargon introduced with English-speaking whale hunting crews especially in Chukotka (de Reuse 1996). English is perhaps the major foreign language in large parts of NEA and there are many native speakers, often soldiers, in Japan and South Korea. Furthermore, it often serves as a *lingua franca* in international communication.

## 2.6 Japonic (Japanese-Ryūkyūan)

The Japonic language family most likely had its origin on the Korean Peninsula and only later expanded into the Japanese archipelago. This expansion is connected with the Yayoi people, originally perhaps farmers along the Yangtze, who after 850 BCE via Korea spread to Japan where they arrived by about 400 BCE (e.g., Janhunen 2003a; Sean & Toshikazu 2011; Hirofumi & Oxenham 2013: 219; Siska et al. 2017: 2f.). The Yayoi people mixed with and replaced the original Jōmon population, their hunter-gatherer lifestyle as well as their languages. Peripheral areas such as Hokkaidō and the Ryūkyūan Islands preserve stronger traces of the Jōmon genome. But while Ainuic languages in Hokkaidō may represent the last remnants of the Jōmon languages, Ryūkyūan languages are clearly related to Japanese. According to Vovin (2013b: 202), the southward migration of Ryūkyūan only started in the 9th century.

According to one classification, Japanese can be divided into Old (592-794), Late Old (794-1192), Middle (1192-1603), and Early Modern Japanese (1603-1867) (Hasegawa 2015: 5ff.). Old Japanese can be further divided into Eastern, Central, and Western Old Japanese. Eastern Old Japanese was spoken in what today is the Kantō area in the 8th century CE, while Western Old Japanese is the language from Nara (Kupchik 2011). Hachijō is the only modern descendant of Eastern Old Japanese (Kupchik 2011: 9). Central Old Japanese, thought to be the predecessor of Modern Japanese, is almost unknown (but see Kupchik 2011: 7f., 852). Old Japanese has to be distinguished from Classical Japanese, which was based on Late Old Japanese as defined above and served as a literary language (Tranter 2012a). There is evidence for the former presence of Para-Japonic or Japonic languages on the Korean Peninsula as well as on Jeju Island (Vovin 2013a), but no information relevant for this study can be obtained from these long-gone varieties (see also Beckwith 2007 and especially Pellard 2005 for some discussion).

Japonic had contact with Ainuic, Koreanic, Sinitic, Amuric, Uilta etc. Modern Japanese, furthermore, has been influenced by several European languages and especially English. Contact with Austronesian on Taiwan led to the emergence of Yilan Creole. The dialects of Japanese as well as Ryūkyūan languages are both increasingly being replaced by Standard Japanese, which itself is based on the Tōkyō dialect in the Eastern dialect area (Sanada & Uemura 2007). Yilan Creole is under Chinese influence.

## 2.7 Koreanic

The internal dialectal differences of Korean should not be underestimated, and some of these dialects, notably Jeju on Jeju island and Yukcin in the Northeast, have been said to exhibit language-like differences with regard to other varieties of Korean. It is therefore possible to speak of the *Koreanic* language family instead of a *Korean* isolate. Regarding the origin of Koreanic, Vovin (2013b: 201) has recently argued for a location in the north:

It appears that the migration of the Korean[ic] speakers to their present location was quite straightforward, from southern Manchuria in the north to the Korean

Peninsula in the south. The linguistic process of Koreanization took several centuries, and it appears that proto-Korean[ic] or pre-Old Korean gradually replaced [Para-]Japonic languages between the 3rd and 8th centuries CE. The central and southern parts of the Korean Peninsula were originally [Para-]Japonic speaking. (my square brackets)

Today, Koreanic is distributed across the entire Korean Peninsula as well as adjacent parts of China, parts of Sakhalin, and even Central Asia. Theoretically, Central Asian Korean (Kolyemal) as spoken in eastern Uzbekistan, for example, is located outside of Northeast Asia. However, given its location very close to Xinjiang and the fact that it preserves several conservative features that were lost in Korea, it will also be included.

Korean is historically attested in several stages that may be called Old Korean, Middle Korean, and Modern Korean, but recent descriptions disagree on how exactly the historical stages of Korean should be classified. Whitman (2015) considers Old Korean to be the language of Unified Silla (668-935 CE), while Nam (2012: 41) argues that the Old Korean period already began in the 5<sup>th</sup> century CE.

We divide Old Korean (OK) into Early, Mid and Late Old Korean (EOK, MOK, LOK). EOK was the Korean of the Three Kingdoms period, roughly from the start of the fifth century until Silla unified the Three Kingdoms in the 660s. MOK was the Korean of the Unified Silla [Sinla] period, from the 660s until the 930s when Koryŏ [Kolye] re-unified the country. LOK was the language of the earlier part of the Koryŏ dynasty from the 930s till the mid-thirteenth century.

The languages that were spoken before or during Unified Silla are only poorly attested. Very likely these languages included Para-Koreanic and Para-Japonic, but no relevant material is available for the purposes of this study, which is why they have been excluded here altogether. Old Korean was followed by Middle Korean, more exactly Early Middle Korean (10<sup>th</sup> to 14<sup>th</sup> centuries) and Late Middle Korean (15<sup>th</sup> and 16<sup>th</sup> centuries), roughly divided by the invention of the Hangul script in 1446 (Sohn 2012).

Koreanic had contact with Southern Tungusic, Japonic, and Sinitic, which forms a very strong ad- and superstrate. Both Japonic and Koreanic derive a large amount of vocabulary from Sinitic. Today, English is an important contact language as well.

## 2.8 (Khitano-)Mongolic

There are a dozen Mongolic languages and all are spoken in Northeast Asia except for Kalmyk (an aberrant dialect of Oirat) and Moghol in Afghanistan (Janhunen 2003e, 2006). Apart from the Mongolic languages proper, there is what has been termed Para-Mongolic (Janhunen 2003c; 2012a), i.e. sister languages of the Proto-Mongolic lineage (e.g., Khitan). All known Para-Mongolic languages are extinct and given the scarce material, Para-Mongolic languages will be excluded from the discussion. The age of the Mongolic language family, i.e. the time of the break-up of the Proto-Mongolic unity, is thought to be only about 800 years (e.g., Janhunen 2012b: 3). If one includes Para-Mongolic, the family

must be much older, but Janhunen's (2012d: 8) estimate of an age of about 1500 to 2500 years before present shows that the details are far from clear. In addition to the modern Mongolic languages there are historical records of older stages, notably so-called Middle Mongol, which "is the technical term for the Mongolic languages recorded in documents during, or immediately after, the time of the Mongol empire(s), in the thirteenth to the early fifteenth centuries." (Rybatzki 2003b: 57) In addition, there is written Mongol, a literary language written with the Uyghur alphabet that has a history of about 800 years and exhibits several archaic features (Janhunen 2003f). The recently partly deciphered Hüis Tolgoi inscription from Mongolia seems to represent a form of early Mongolic and is considerably older than Middle Mongol (e.g., Vovin 2017). The "homeland" problem is notoriously difficult for many language families. However, for Mongolic it quite clearly was located somewhere in present-day northeastern Mongolia, the place where the Mongolic expansion had its starting point (Janhunen 2003e: xxxiv). But Proto-Mongolic itself formed a larger family with Para-Mongolic, and the question about the original location of this proto-language of Proto- and Para-Mongolic (Janhunen 2012a: 114 proposes the name *Khitano-Mongolic*, also adopted here, and Shimunek 2014; 2017 *Serbi-Mongolic*), is less easy to answer. Janhunen (2012d: 10) assumes that it was located further to the south in present-day Liaoning or eastern Inner Mongolia:

There is a particularly clear parallelism in the expansion of the Mongolic [including Para-Mongolic] and Tungusic language families. Once they had occupied their protohistorical positions on both sides of the Liao basin, they both assumed a general northward trend of expansion. In the light of the available information on the history and protohistory of the region, the Mongolic homeland has to be placed in southwestern Manchuria (Liaoxi), while the Tungusic Homeland can hardly have been located anywhere else but in southeastern Manchuria (Liaodong), though quite possibly also extending to the northern part of the Korean Peninsula. (my square brackets)

On Tungusic, see §2.10. Janhunen's assumption of a Pre-Proto-Mongolic homeland situated roughly in eastern Manchuria is corroborated by some historical facts, such as the Khitan Liao-dynasty (᠊ᠯᠢᠠᠤ, 916-1125 CE) that roughly derived from this region.

Mongolic in general shows strong influence from Turkic languages and *vice versa* (Schönig 2003). Individual Mongolic languages participated in different linguistic areas that sometimes overlap and display a different strength of convergence. Shirongolic is an integral part of the so-called Amdo Sprachbund. Dagur, together with the two Tungusic languages Solon and Oroqen, formed a small linguistic area for itself, but during the Qing-dynasty (1636-1911) were also under the strong influence of yet another Tungusic language, Manchu. Similar to Tungusic, Mongolic languages today can be classified as to whether they are under the influence of the national language of Russia (Kalmyk, Buryat) or China (Dagur, Shirongolic etc.). But unlike Tungusic, this only partly applies to the Mongolic languages spoken in "Outer Mongolia", where Russian influence appears to be receding, and does not apply at all to Moghol in Afghanistan. A national language itself, Mongolian of course influences all Mongolic languages spoken in Mongolia.

## 2.9 Trans-Himalayan (Sino-Tibetan)

It has been pointed out that the name *Sino-Tibetan* is somewhat misleading and it will not be used in this book. The traditional view, as advocated by LaPolla (2013), for example, claims that Sino-Tibetan has two main branches, Sinitic and Tibeto-Burman. According to this view, the origin of Sino-Tibetan (and not only of Sinitic) is usually said to have been around the Yellow River. Some of the justified criticism to previous approaches to the family has been aptly summarized by Blench & Post (2014: 93):

“Reconstructions” have been proposed which have failed to take many languages of high phyletic significance into account; these forms have been repeatedly quoted without remark in the literature, in the process gaining a lustre they hardly deserve. Sino-Tibetan has no agreed internal structure, and yet its advocates have been happy to propose dates for its origin, expansion and homeland in stark contradiction to the known archaeological evidence. A focus on “high cultures” (Chinese, Tibetan, Burmese) has led to an emphasis on these languages and their written records, something wholly inappropriate for a phylum where an overwhelming proportion of its members speak unwritten languages.

Therefore, the more adequate and neutral name *Trans-Himalayan* (van Driem 2014) will be employed here instead, which does not imply a split into only two main branches and suggests an origin and center of diversity further to the southwest. In fact, most Trans-Himalayan languages are located in South or Southeast Asia. According to van Driem (2014) and Blench & Post (2014), the geographical distribution of the different branches suggests an origin of the whole language family in the eastern Himalayas. Under this assumption, Sinitic would be the northernmost of many different branches of the family. Needless to say, this innovative view is not yet accepted by all researchers and deserves further investigation (see LaPolla 2016 for a discussion).

This study only includes languages from three of a total of perhaps 42 different sub-branches of Trans-Himalayan (van Driem 2014), namely Sinitic, Tibetic (a subbranch of Bodish), and Qiangic. The age of *Sinitic* depends on the definition. Traditionally, old stages of Chinese are divided into *Old Chinese* and *Middle Chinese*. However, a new approach developed by Norman (2014), which focuses on evidence from the spoken languages, makes a distinction into *Common Dialectal Chinese* (CDC, the proto-language of all modern Chinese languages except Min) and *Early Chinese* (EC, the proto-language of Min, CDC etc.). Roughly speaking, CDC can be compared with the Romance languages and Early Chinese with Italic. If *Sinitic* refers to CDC and its descendants, then the age is perhaps about 2000 years. If, however, *Sinitic* refers to the whole branch of Trans-Himalayan (i.e., (pre-)EC), then Sinitic is perhaps some 1500 years older. The latter view will be adopted here. However, Norman was reluctant to estimate the ages of the two proto-languages. While Norman’s is perhaps the best approach to the history of Chinese yet, this study necessarily takes a pragmatic stance. Compared with Indo-European, the reconstruction of Chinese is still in its infancy and goes beyond the possibilities of this study, which will mostly be focusing on modern Chinese languages. In order to capture some of the history of Chinese, I will refer to the recent study by Baxter & Sagart

(2014a,b), who employ the term *Old Chinese* as a more or less useful cover term for the earlier period of Sinitic:

We use the term “Old Chinese” in a broad sense to refer to varieties of Chinese used before the unification of China under the Qín 秦 dynasty in 221 BCE. The earliest written records in Chinese are oracular inscriptions on bones and shells from about 1250 BCE (in the late Shāng 商 dynasty, which was overthrown by the Zhōu 周 in 1045 BCE), so this is an interval of about 1,000 years. Obviously there must have been many varieties of Chinese during this period, widely distributed in time and space. (Baxter & Sagart 2014a: 1)

Throughout its history, Sinitic had intense language contacts with many surrounding languages (see Matthews 2010). Especially intense was the influence on Korean and Japanese, which derive a large amount of their vocabulary from Sinitic. Mandarin today is the dominant language of China, and has already started to replace several minority languages throughout the country. Just like Russian dominates the northern half of NEA, Mandarin has a leading position in the southern half.

Following Tournadre (2014), it is perhaps best not to speak of the Tibetan, but of the Tibetic, branch, which goes back to Old Tibetan (ca. 7th to 9th century CE) as its proto-language, which is closely related to the Classical Tibetan language:

‘Classical Tibetan’ is an idealization, referring both to over a millennium of written history and to a tradition of prescriptive grammar which many of the authors of the texts, in some cases down to the present, made greater or lesser efforts to conform to. [...] The term ‘Old Tibetan’ is used to refer to written material from before about 1000 CE, primarily inscriptions and documents found in the Dun-huang caves (DeLancey 2003: 255f.)

Today Tibetic encompasses about 200 different varieties distributed over an extremely large area, which can, according to Tournadre (2005), be classified into eight “sections”. Only some varieties from the eastern (Baima, Cone, Zhongu) and northeastern sections (Amdo Tibetan, gSerpa) will be included here. Amdo Tibetan is of special importance for this study because of its dominant position in the Amdo Sprachbund (Sandman & Simon 2016, §3.5).

Whether Qiangic is a valid subgroup of Trans-Himalayan, and which languages it should cover, is an ongoing debate. Chirkova (2012) argues that it should be reconceptualized as an areal rather than a genetic group of Trans-Himalayan languages. Without a final solution to the problem at hand, this study retains the common designation as Qiangic, which is first and foremost a pragmatic decision. In NEA only one language is usually classified as Qiangic:

Tangut (also known as the Xixia language) is an extinct Tibeto-Burman language that was spoken in the Xixia empire that existed from 1038 to 1227 in northwestern China. The language was buried in oblivion till 1908 when the Russian geographer P.K. Kozlov discovered the ruins of a Tangut city at Khara Khoto. (Gong Hwang-Cherng 2003: 602)

Baima, tentatively classified as Tibetic here, is sometimes also treated as a Qiangic language (Chirkova 2012: 139).

There have been many attempts to connect Trans-Himalayan with other language families, none of which is widely accepted. A *Sino-Tibetan-Austronesian* hypothesis that also includes Tai-Kadai as a branch of Austronesian is currently being debated (see Sagart 2016), but does not seem to be gaining acceptance.

## 2.10 Tungusic

Tungusic is the name of a language family that includes about a dozen to twenty different languages distributed over a vast area in Siberia and Northern China. Experts do not agree on the exact number of languages, primarily because of the complex network of dialects and mutual influence. Instead of *Tungusic*, some researchers prefer the name *Manchu-Tungusic* (e.g., Pevnov 2012), but I will continue to use the name *Tungusic* as a convenient label for the whole language family. The name *Tungusic* historically referred to the Evenki or the Even and their languages, but today does not designate any specific variety. In addition, if understood in the old sense, the name *Manchu-Tungusic* actually refers to only two or three of many more languages. In addition, the term suggests a primary split of the language family into Manchu and Tungusic, which is not necessarily accurate (e.g., Ikegami 1974; Georg 2004; Janhunen 2012d; Hölzl 2015a; 2017a). What is more, the name *Tungusic* belongs to a long tradition of referring to the whole language family (e.g., Benzing 1956).

Tungusic today is usually classified into four different groups (Ikegami 1974; Georg 2004), which can be called Jurchenic, Nanaic, Udegheic, and Ewenic (Janhunen 2012d). According to one hypothesis that will be followed here, the first two form the southern Tungusic branch, and the latter two the northern branch. Janhunen (1996; 2005; 2012d) assumes that Proto-Tungusic was spoken in southern Manchuria, east of the Liao river and partly in the north of Korea:

The linguistic facts suggest that the Tungusic family represents a classic case of language spread from a relatively compact homeland. Against the overall ethno-historical picture of Northeast Asia, it appears likely that the Tungusic homeland was located in the region comprising Southern Manchuria and Northern Korea, the historical habitat of the Jurchen-Manchu. From here Tungusic expansion took Tungusic to the Amur basin, where Nanai, Udeghe, and Ewenki branches subsequently emerged. These initial expansions of Tungusic may have taken place between 2000 and 1000 years ago (Janhunen 1996: 216-233). (Janhunen 2005: 39)

But a more plausible location appears to have been further north, as has also been claimed by Pevnov (2012) and Vovin (2013b). An educated guess for an original location of Tungusic should probably pinpoint the confluence of the Amur, the Sunggari, and the Ussuri. From this region Jurchenic expanded southwards along the Sunggari and the Ussuri, Nanaic followed the lower Amur northwards, Udegheic spread along the eastern

tributaries of the lower Amur and the Ussuri, and Ewenic speakers migrated along the Amur river towards the northeast and to some extent followed the left tributaries such as the Bureya and the Zeya. Parts of Ewenic (mostly Evenki and Even) then rapidly covered almost all of Siberia. This expansion of Ewenic has also been recognized by Janhunen (2005: 39):

The modern distribution of Tungusic is largely the result of the secondary expansion of the Ewenki branch, which very probably began from the Middle Amur region no more than 1000 years ago. This expansion spread Tungusic over the whole of Siberia, from the Okhotsk Sea in the east to the Yenisei basin in the west, and from Lake Baikal in the south to the Arctic Ocean in the north. The expansion has continued until recent times, especially in Northeast Siberia. Territories reached only in the 19th century include Kamchatka (Ewen) and Sakhalin (Siberian Ewenki).

Janhunen (2005) is right in pointing out the internal homogeneity of both Evenki and Even, which indicates a very recent spread. Even today the number of Ewenic languages is highest in Manchuria.

A recent study found evidence that the direct ancestors of some Tungusic-speaking peoples have been living in Manchuria for at least 8000 years:

We report genome-wide data from two hunter-gatherers from Devil's Gate, an early Neolithic cave site (dated to ~7.7 thousand years ago) located in East Asia, on the border between Russia and Korea. Both of these individuals are genetically most similar to geographically close modern populations from the Amur Basin, all speaking Tungusic languages, and, in particular, to the Ulchi. (Siska et al. 2017: 1)

This is no proof, of course, that the ancestors of the Tungusic language family were spoken in the area as well. However, the genetic continuity might suggest that there may not have been a language shift from some unknown languages to the Tungusic languages family (or its predecessor), which would be expected to leave clearer traces of genetic admixture. Another recent genetic study, for example, found that the Udihe appear to be “the result of admixture between local Amur-Ussuri populations and Tungusic populations from the north.” (Duggan et al. 2013: 1) Unfortunately, it is still too early to draw any substantial linguistic conclusions based on these results.

There are too many instances of language contact of Tungusic languages all over NEA to be summarized here in detail. Manchu used to be an important superstrate language for all languages in Manchuria and also had a certain impact on Mandarin (e.g., Tsumagari 1997). Manchu itself has a pronounced Mongolic, Para-Mongolic, and perhaps Koreanic adstrate. Sibe had contact with Mongolic languages such as Khorchin and, the group of speakers who were relocated to Xinjiang in 1764, with several Turkic languages such as Uyghur. Several Tungusic languages had contact with Amuric languages along the lower Amur. Evenki had contact with several Mongolic languages such as Buryat, with Nivkh on Sakhalin, as well as with Yakut, Yeniseic, Yukaghiric, and Samoyedic. Even had

contact with Chukotko-Kamchatkan as well as Yakut, and partly replaced Yukaghiric. Oroqen and especially Solon had an almost symbiotic relation to the Mongolic language Dagur (e.g., Janhunen 1997). The same is true for the two Khamnigan Evenki dialects with Khamnigan Mongol (e.g., Janhunen 1991; Janhunen 2003b). Before the advent of Russian and Mandarin influence, Khitano-Mongolic exerted the most important influence over all of Tungusic (Doerfer 1985).

## 2.11 Turkic

Turkic languages are widespread today, from the Arctic Sea in the north to Qinghai in the south and from Manchuria in the east to Turkey in the west (excluding recent migrations to Germany, for instance). The spread of Turkic all over Eurasia had its beginnings in southern Siberia and northern Mongolia, where the oldest Turkic records, the Orkhon inscriptions, were found (Golden 1998). Turkic has perhaps six main branches, Oghur, Khalaj, Siberian, Uyghur-Karlak, Kipchak, and Oghuz (Johanson 1998: 81f.; Johanson 2006a: 161f.). First Oghur, today only represented by Chuvash in European Russia, and then perhaps Khalaj (in Iran) split away from the rest. Most languages covered here are from the Siberian branch, but languages from all branches except Oghur and Khalaj are today located in NEA. This study excludes the by now perhaps extinct archaic Turkic language Khotong from Mongolia, for which no data are available to me (Shimunek et al. 2015: 148).

The classification above only includes modern Turkic languages, but there are historically attested varieties of Turkic that will be briefly mentioned as well, notably Old Turkic and Chaghatay.

Old Turkic is taken to be the language underlying three corpora. The first one consists of official or private inscriptions in the runiform script, dating from the seventh to tenth centuries, in the territory of the second Türk empire and the Uyghur steppe empire - present-day Mongolia - and the Yenisey basin. The second and most extensive corpus consists of ninth to thirteenth century Old Uyghur manuscripts from northwest China in Uyghur, runiform and other scripts. [...] The third corpus consists of eleventh-century texts from the Karakhanid state, mostly in Arabic script [...]. (Erdal 1998: 138)

Chaghatay can be defined as a succession of stages of written Turkic in Central Asia. In many respects it is also a continuation of earlier stages, notably of Karakhanid Turkic, with Kharezmian Turkic as a transitional stage. It cannot be defined as a fixed entity in time and space. Chaghatay sources are a hybrid collection of different varieties of Turkic, who from the late fifteenth century onwards more or less tried to focus on a specific model known as Classical Chaghatay. (Boeschoten & Vandamme 1998: 166)

Chaghatay influenced several written languages, including the Kipchak languages Tatar and Kazakh, the Oghuz language Turkmen, and the Uyghur-Karluk languages Uzbek and

Uyghur (Boeschoten & Vandamme 1998: 168). In fact, the Uygur-Karluk branch is sometimes also called the Chagatay branch of Turkic.

The extensive contact between Turkic and other languages has been summarized by Schönig (2003) and Johanson (2010). Turkic languages in general had strong contact with Mongolic. But individual languages underwent a plethora of contact situations that cannot all be summarized here. Yakut had contact with Buryat, and later with Evenki, Yukaghir, and Nganasan, which led to the emergence of Dolgan. In the southwest there is contact with Iranian and in the southeast with Sinitic. In the Amdo region there is a strong interaction with Mongolic and Sinitic varieties as well as with Amdo Tibetan.

## 2.12 Uralic

Uralic (e.g., Sinor 1988; Abondolo 1998) is a language family with a very long history comparable to that of Indo-European. The primary split of the language family separates the Samoyedic languages from Finno-Ugric. Despite the rather small comparative corpus between Finno-Ugric and Samoyedic, their genetic relation is usually recognized. Proto-Samoyedic perhaps split about 2000 years ago, while Finno-Ugric and Samoyedic had a common origin in Proto-Uralic about 5000 years ago (Janhunen 2009: 68). The location of the Uralic homeland is disputed, but Janhunen (2009: 71) argues for “the borderline between the Ob and Yenisei drainage areas in Siberia” and thus for a region at the edge of NEA. Given the connection of Uralic with Yukaghiric (§2.14), Pre-Proto-Uralic could even have been spoken in NEA. However, only the Samoyedic branch is clearly represented in NEA (e.g., Janhunen 1998).

Listed roughly from north to south, these are (older designations given in parentheses): Nganasan (Tavgy), Enets (Yenisei-Samoyed), Nenets (Yurak), Selkup (Ostyak-Samoyed), Kamass(ian), and Mator (Motor). The southernmost languages Kamass and Mator, are now no longer spoken: Mator was replaced by Turkic idioms during the first half of the nineteenth century, and the fact that it is known at all today is because of intensive philological work done with word lists; the last Kamass speaker died in 1989. [...] only Nenets is spoken by a relatively large number of people (some 27,000); Selkup, which has sharp dialectal divisions, has fewer than 2,000 speakers; Nganasan, some 600; and Enets, perhaps 100. (Abondolo 1998: 2)

Elena Skribnik (p.c. 2017) informed me that in NEA there are also a few speakers of, for example, Estonian. However, such isolated groups will mostly be neglected in this study (but see Miestamo 2011 and §5.12.2). Together with Yeniseic, Samoyedic forms the western border of the NEA area. Samoyedic may have been spoken by the Tagar culture (ca. 1000-200 BCE) in the Minusinsk basin (Janhunen 2009: 72; Parpola 2012: 294), but this remains somewhat speculative. Just like Yeniseic, Samoyedic spread along the Yenisei northwards, while those varieties left behind were slowly replaced by languages from other families.

Samoyedic had contact with several Finno-Ugric, Yeniseic, and Turkic languages as well as Evenki, Russian, and perhaps some early form of Tocharian. Selkup had an especially strong interaction with the Yeniseic language Ket.

## 2.13 Yeniseic (Yeniseian)

Typologically, Yeniseic is the most atypical Siberian language family (§3.5). Today it is represented by only one language, namely Ket. But there used to be several other Yeniseic languages (Arin, Assan, Kott, Pumpokol, Yugh) that have since disappeared. Yeniseic substrate toponyms, largely river names that have endings such as *-ul*, *-ses*, or *-det*, cover a large area from the Irtysh in the west to northern Mongolia in the east and indicate a more widespread distribution in the past (Vajda 2009a: 474). The homeland of Yeniseic may have been the Altai region, especially the Karasuk culture (1200-700 BCE) (Flegontov et al. 2016: 1f.). According to Vajda (2010: 33), less than 100 Ket are still able to speak the language.

There is some evidence to suggest that a Yeniseic language was one of the language of the historic Xiongnu (匈奴) in northern China, the main rivals of the Han dynasty (206 BCE to 220 CE) (cf. Vovin et al. 2016 and references therein). In addition, Vajda (2010) has made a strong argument for a genetic connection between Yeniseic and Na-Dene languages, called the *Dene-Yeniseian hypothesis*. Apart from Eskaleut, this would be the first language family discovered that connects languages in Asia and the Americas. The theory is currently gaining acceptance as new pieces are added to the puzzle (e.g., Vajda 2013), and, at least for the moment, it seems that there are fewer critics than proponents. Nevertheless, more research over the following years will show whether the hypothesis can stand the test of time. If Dene-Yeniseian turns out to be a valid genetic unit, there are several different possible explanations for its modern distribution. One possibility would be to assume a location of the proto-language somewhere in (south)eastern NEA. From there, Yeniseic moved westwards, whereas Na-Dene moved northwards to finally cross Beringia. But Sicoli & Holton (2014) have recently argued for an alternative that assumes an original location in the Beringian area. Yeniseic, according to them, is the result of a back-migration into Asia. However, this goes against the general rule of thumb that migrations in NEA usually follow a south-to-north direction. In any case, the migration of Yeniseic down the Yenisei and of Na-Dene from Alaska southwards are widely accepted and must be common ground for any additional hypothesis. The question of the time depth of the hypothetical Proto-Dene-Yeniseian language remains unsettled for now, but must necessarily be many thousand years older than Proto-Yeniseic (see §5.13.4).

## 2.14 Yukaghiric

The term *Yukaghiric* is employed here to refer to the language family usually called *Yukaghir*. However, there are two rather different extant Yukaghiric languages, which is why a specialized designation for the language family seems in order to avoid confusion.

These two languages are called Kolyma Yukaghir (Odul) and Tundra Yukaghir (Wadul). Tsumagari et al. (2007) classify Tundra Yukaghir as “seriously endangered” and Kolyma Yukaghir as “moribund” as there are only several dozen elderly speakers left for both languages (Matić 2014: 130). Yukaghiric languages were mostly replaced by Even (Tungusic), Yakut (Turkic), Chukchi, and Kerek (Chukotko-Kamchatkan), as well as Russian (Slavic). War with and exploitation by the Russians, together with smallpox epidemics, decimated their number drastically. Where there were an estimated 4500-5000 Yukaghir in the 17<sup>th</sup> century, only 150-200 remained at the end of the 19<sup>th</sup> century, but their number has been growing again ever since (Rédei 1999: 3; Forsyth 1992: 74-80).

Yukaghiric languages must have been extremely widespread in northeastern Siberia until the 17<sup>th</sup> century. According to Volodko et al. (2008), the Yukaghir were even involved in the formation of the Samoyedic-speaking Nganasan much further to the west. Even so, they seem to have reached the northern parts of NEA from a location further south. Häkkinen (2012: 93) argues that

Yukaghir[ic] can be derived from the west, as it was spoken earlier near the Lena. We may assume that Yukaghir[ic] at some point in the past migrated down the Lena, just as Yakut did later, and that Early Proto-Yukaghir[ic] was spoken somewhere near the Upper Lena and the region of Lake Baikal, the watershed area between the Lena and Yenisei river systems. (my square brackets)

If Häkkinen’s assumption is correct, this brings Yukaghiric geographically much closer to other language families such as Tungusic, Samoyedic, Khitano-Mongolic, Turkic, and Yeniseic. A southern origin of the Yukaghir is also corroborated by evidence from mitochondrial DNA analyses (Volodko et al. 2008). Häkkinen’s conclusions are built on an assumption of a direct contact of Yukaghir with Uralic languages. Janhunen (2009: 61) explicitly denies a connection between Uralic and Yukaghiric. But most researchers do not exclude the possibility of a genetic connection (e.g., Pispane 2013) or at least contact (e.g., Rédei 1999; Aikio 2014). The separation of the two Yukaghiric languages has been estimated to date back to about 2,000 years ago (Maslova 2003a: 28), which remains rather speculative and might be an overestimation. For instance, personal pronouns in the two extant Yukaghiric varieties are basically identical, which would not be expected after such a long period of separation. The location of Pre-Proto-Yukaghiric in the south of NEA, on the other hand, must be much older and has been tentatively dated to the early-middle Holocene (Volodko et al. 2008: 1097) and thus might be much earlier than “Early Proto-Yukaghir” as was assumed by Häkkinen (2012: 93).



## 3 Areal typology and Northeast Asia

Chapter 2 has introduced the languages of Northeast Asia from a genetic perspective, i.e. classified into language families. The focus in the present chapter is on language contact instead and adds an areal perspective to the discussion. The two classifications are not always clearly separable, especially at greater time depths (e.g., Nichols 2010; Operstein 2015), and given the fact that, naturally, languages from one family can also have contact with each other (e.g., Epps et al. 2013). Since an exhaustive presentation of all language contact phenomena goes well beyond the possibilities of this study—“Language contact is everywhere” (Thomason 2001: 8)—, there will be a focus on some points that are especially relevant.

### 3.1 Theoretical considerations

This chapter is concerned with structural diversity, or rather, structural similarity among languages. There are several different reasons that languages can be similar, including universals, tendencies, chance, genetic inheritance, and language contact (e.g., Aikhenvald & Dixon 2001: 1-3). It seems that all languages around the globe have specialized constructions for asking questions, so that this is a linguistic universal and the reason this study is possible in the first place. Interestingly, there might even be universal questions such as ‘What is your name?’, ‘Who are you?’, and ‘What is that?’ that are, however, expressed differently from language to language.<sup>1</sup> There may be yet more specific universals. Dingemanse et al. (2013: 1) have quite convincingly shown that the repair initiator *huh?* could well be a universal word “not because it is innate but because it is shaped by selective pressures in an [enchronic] interactional environment that all languages share: that of other-initiated repair.” In my opinion, potential universals of this kind have to be distinguished from strong tendencies, such as the fact that positive one-word answers in a great many languages around the globe contain laryngeal sounds [h] and [ʔ] (Parker 2006). Take German, for instance, which has the word *ja* ‘yes’. At first glance, this does not contain any laryngeal sounds, but it has many different variations, among which one encounters [jaʔ] with a final glottal stop as well as ingressive [hja↓] with an initial laryngeal fricative (my knowledge). A similar tendency is for languages to have rising intonation in polar questions, which is common but by no means universal. Hawai’i Creole English, for example, has falling intonation instead (Veluppilai 2012: 353). A factor that should not be underestimated is chance resemblance. An example from the category

---

<sup>1</sup>David Gil (p.c. 2018) informs me that he is working on a typology of the question ‘What is your name?’ on which see also Idiatov (2007); Hölzl (2014b) and §§4.3.1, 5.6.3.

of questions are the polar question markers  $-(V)?$  in Hup (Nadahup, Epps 2008: 784ff.) in South America and  $-?$  in Crow (Siouan, Graczyk 2007: 391) in North America that at least in some instances are basically identical. To my knowledge there does not appear to be a general tendency for question markers to exhibit laryngeal sounds, as far as we know the two languages do not share a common ancestor, and there certainly was no contact between them. This only leaves pure coincidence to account for this similarity. An example for a chance resemblance in the interrogative system would be Tocharian (Indo-European) *kos* and Dolgan (Turkic) *kas* ‘how much’. Given that both interrogatives and question markers tend to be very short, chance resemblance is extremely hard to distinguish from genetic inheritance and language contact. As seen in Chapter 2, genetic inheritance refers to languages of one and the same language family that go back to one proto-language and therefore preserve features that are similar to each other. The two Tungusic languages Evenki and Even, to take a random example, shared a common ancestor only several centuries ago and therefore display many similarities such as an almost identical question marker  $=Ku$  (§5.10.2). The last of the explanations for the similarity between different languages is language contact. The term *language contact*, of course, is nothing but a metaphorical abstraction of what is actually an integral part of the complex interaction of different human beings. But certainly it serves its purpose to facilitate our discourse on the topic. Language contact presupposes a linguistic interaction of speakers of different languages (Thomason 2001: 1f.). Perhaps every linguistic interaction has certain properties that qualify as language contact. However, language contact is usually identified through observable results such as the borrowing of elements. Contact may be either direct or indirect. The latter can be further divided into the contact of two languages with a transmitting language on the one hand, and a common contact language of two languages on the other.

The outcome of language contact differs from instance to instance. Thomason (2001: 10) suggests “a hierarchical set of typologies, starting with a three-way division at the top level into contact-induced language change, extreme language mixture (resulting in pidgins, creoles, and bilingual mixed languages), and language death.” In NEA there are examples of all three kinds, but the details of Thomason’s (2001: 60) typology are too complex to be repeated here in full. Language shift, which today is an extremely common phenomenon around the globe and in NEA, will for the most part be excluded for the lack of relevant data concerning the effects on the grammars of questions. Examples for extreme cases of language contact found in NEA include two extinct pidgins (Chinese Pidgin Russian, Govorka, §5.5), some creolized languages (e.g., Gangou, Wutun, Tangwang, Hezhou, §5.9), several mixed languages (Eynu, §5.11, Mednyj Aleut, §5.4, some Tungusic languages §5.10, §6.3, and an Ainu-Itelmen hybrid), and perhaps some slightly less extreme cases such as Mandarin or Manchu (e.g., McWhorter 2007).

Contact-induced change has several subtypes (*relabeling*, *calquing* etc.), but arguably, regarding the grammar of questions, the most important case is *borrowing*, simply put the transfer of a certain element from one language to another. But how do we actually know that a linguistic item in a given language can be explained by language contact rather than genetic inheritance? Let me illustrate this with an example from the Tundu-

sic language Uilta (§5.10.2). Uilta has a content question marker =*ga* ~ =*ka*. A comparison with closely related languages such as Nanai shows that the marker is not present, in fact, content questions in Nanai remain unmarked. The fact that related languages do not show this marker in most cases rules out an explanation in terms of genetic inheritance. Then the form might simply be an innovation found in Uilta, but no plausible etymology is known to me. Thus, Uilta perhaps borrowed the question marker from a surrounding language. Uilta is spoken on Sakhalin where it is known to have had contact with the neighboring language Nivkh (Yamada 2010, §5.2.2). In fact, Nivkh has overt content question markers, one of which has the form =*ŋa*. Of course, Uilta could also have borrowed the question marker from other surrounding language families such as Japonic. Tsuken, for example, has a content question marker =*ga*. However, Tsuken is spoken in the Ryūkyūan Islands several thousand miles south of Sakhalin. This geographical distance makes a connection extremely implausible, because the speakers of Uilta and Tsuken quite certainly had no direct contact with each other. But what about Japanese, which was once spoken on Sakhalin and has a question marker *ka* か that can also be found in content questions? First of all, Uilta had much more longstanding and intimate contact with Nivkh than with Japanese. However, in order to refute this possibility, more information on Japanese and Uilta is in order. Old Japanese already possessed the question marker, which had more or less the same form, but in Uilta there are further forms such as =*ge* (alternatively written with a schwa ə and an optional long vowel). Given that Uilta has vowel harmony in which *a* stands opposed to *e* (Tsumagari 2009b: 3), this appears to be an innovation and the form might still derive from either Nivkh or Japanese. However, the integration of the question marker into the morphological system suggests a relatively early borrowing, which makes a comparison with Japanese much less likely. Furthermore, Nakanome (1928: 50ff.) mentions a form that was written as <ñö>. The pronunciation of this form must be [ŋə], as a comparison of Nakanome (1928) with Ikegami's (1997) modern dictionary suggests, e.g. <önnö> = [ənnə] 'mother', <ñâla> = [ŋaala] 'hand'. The existence of the velar nasal makes a comparison with Nivkh much more likely than with Japanese. The fact that both Nivkh and Uilta, but not the surrounding languages, overtly mark polar and content questions differently—i.e. there is a similarity in type—confirms this hypothesis (e.g., Hölzl 2015e). This typological parallel has also recently been observed by Pevnov (2016: 59f.). On the contrary, Japanese allows the marker *ka* in both polar and content questions. For reasons of space, this procedure will not be given in full detail for every potential instance of borrowing identified in Chapter 5. A list of all the borrowed elements of the grammars of questions in NEA found throughout this study is given in Chapter 6. In several cases the details will have to be discussed by experts of the individual languages.

One of the central concepts of areal linguistics is the heavily disputed notion of a *linguistic area* or *sprachbund*. The best summary of previous approaches can be found in Campbell (2006: 18), whose rather skeptical conclusion is the following.

Every 'linguistic area', to the extent that the notion has any meaning at all, arises from an accumulation of individual cases of 'localized diffusion'; it is the investigation of these specific instances of diffusion, and not the pursuit of defining prop-

### 3 Areal typology and Northeast Asia

erties for linguistic areas, that will increase our understanding and will explain historical facts.

There is a strange dissonance between theoretical approaches that usually take a negative stance on the concept (e.g., Dahl 2001; Bisang 2010) and the widespread use of the term for individual areas such as the Amdo Sprachbund. This study acknowledges the fundamental theoretical problems of the concept, but takes a pragmatic approach. The term *linguistic area* is taken as a useful label if it is not meant to indicate clear-cut boundaries or absolute homogeneity. Like many linguistic phenomena, linguistic convergence is obviously a matter of degree (cf. Langacker 2008: 13) and there is no problem in calling areas of strong convergence a *sprachbund* or *linguistic area*. As a rule of thumb, an area should be characterized with the help of features that are not very common cross-linguistically, and that are not shared with surrounding areas. NEA is surrounded by several possible areas such as the *Greater Himalayan Region* to the south (Kraaijenbrink et al. 2009) and the *Pamir-Hindukush Sprachbund* (Novák 2014: 82) as well as the *Araxes-Iran Linguistic Area* to the southwest (Stilo 2015), and *Mainland Southeast Asia* (Enfield & Comrie 2015) to the southeast. Unfortunately, with only the exception of MSEA (Enfield & Comrie 2015), the definition of all of these areas is quite problematic. Nevertheless, the fact that the entire southern and southeastern boundary is marked by mountains teeming with linguistic diversity indicates that they form an *accretion* or *residual zone* (Nichols 1992; 1997; 2015, see §3.4) that functions as some kind of boundary. The most difficult problem is the identification of a western boundary (Heggarty & Renfrew 2014a: 873). Immediately to the west of NEA live the speakers of the Uralic, more precisely Finno-Ugric, languages Khanty and Mansi that are sometimes collectively called Ob-Ugric. Their genetic classification is disputed, with some arguing that they belong to a single branch and others for a classification into two different branches that had strong mutual contacts, called Khantic and Mansic by Janhunen (2009: 65). It is difficult to consider these two languages as forming a useful western boundary. But the Western Siberian Lowland together with Kazakhstan to its south is a region of low linguistic diversity (a spread zone, Nichols 1992), which contrasts with the adjacent areas of NEA along the Yenisei. Located to the west of the Ural mountains, and thus separated from NEA by the Western Siberian Lowland, lies the *Volga-Kama Area* (see Manzelli 2015). This is an area of strong linguistic convergence between several Finno-Ugric and Turkic languages (see §5.11, §5.12). If one was to extend NEA to include all of the area to the east of the Ural mountains into, say *Northern Asia* (Nichols 1992: 25f.), the Volga-Kama area would certainly function as a better western boundary than does Ob-Ugric. Nevertheless, several languages with affinities to NEA, notably Finnish or Turkish would still be located to the west of the Volga-Kama Area.

For practical purposes, Eurasia will be treated as a macro-area (§3.2) that contrasts relatively sharply with Mainland Southeast Asia (§3.3) and contains a meso-area called Northeast Asia (§3.4). NEA in turn encompasses several possible micro-areas such as the so-called *Amdo Sprachbund* (§3.5) only some of which will be mentioned in this chapter.

### 3.2 The Eurasian macro-area

As is well-known, NEA is part of a large Eurasian area that is characterized by several dominant features (Table 3.1). This area includes most of Eurasia but not Mainland Southeast Asia (MSEA), parts of Europe and parts of the Near East. There is some variation in the geographical distribution of these features, but notably NEA invariably shares all of them. The variation concerns the periphery of the Eurasian area such as Europe.

The usefulness of some of these word order features for the identification of linguistic convergence is somewhat reduced by the existence of implicational hierarchies connecting several of them (e.g., Bisang 2010: 422; Dryer 2013a). Nevertheless, they define a relatively clear-cut boundary towards the southeast. A possible further trait of this Eurasian Area is the existence of K-interrogatives (§6.2.1).

### 3.3 Mainland Southeast Asia

The sharpest contrast of NEA with other areas is that with Mainland Southeast Asia (MSEA), the adjacent region to the southeast, which has recently been defined as

the area occupied by present day Cambodia, Laos, Peninsular Malaysia, Thailand, Myanmar, and Vietnam, along with areas of China south of the Yangtze River. Also sometimes included are the seven states of Northeast India, and—although here the term ‘mainland’ no longer applies—the islands from Indonesia and Malaysia running southeast to Australia and West Papua (Enfield & Comrie 2015: 1)

MSEA is widely accepted as a region of strong convergence of five different language families, namely Trans-Himalayan (Sino-Tibetan), Tai-Kadai, Hmong-Mien (Miao-Yao), Austroasiatic, and Austronesian. My definition of NEA excludes the Yangtze watershed, a part of which is likely the historical homeland of the Hmong-Mien languages (Ratliff 2010: 241) that clearly belong to the MSEA area. Moreover, Sinitic languages show an internal split between northern and southern varieties (e.g., Ramsey 1987: 19-26; Matthews 2010: 760f.). In a certain sense, the distinction between Mandarin and Southern Sinitic varieties is symptomatic for the difference between NEA and MSEA. Mandarin is rather homogeneous and is spread over a vast area ranging from Yunnan in the Southwest to Heilongjiang in the Northeast and from Jiangsu in the east to Xinjiang in the west. Southern Sinitic, on the other hand, is limited to a much smaller geographical area but nevertheless shows extremely strong internal variation with many mutually unintelligible varieties (Kurpaska 2010).

There is a qualitative difference between these two areas. The Mandarin area, on the one hand, is unusually uniform; virtually all of the dialects spoken there are mutually intelligible—or very nearly so. [...] But the non-Mandarin area is extremely varied, and within it sharply divergent forms of speech are often separated by only a few miles. (Ramsey 1987: 21)

### 3 Areal typology and Northeast Asia

Table 3.1: Some areally dominant features of NEA shared with neighboring areas

Feature	Approximate area	Source
SOV (SV & OV)	Eurasia, excluding MSEA and Europe	Dryer 2013i
noun-adposition (NAdp)	Eurasia, excluding MSEA, most of Europe, and parts of the Near East	Dryer 2013c
genitive-noun (GenN)	Eurasia, excluding southern MSEA, western Europe, and the Near East	Dryer 2013f
adjective-noun (AdjN)	Eurasia, excluding MSEA, western Europe, and the Near East	Dryer 2013b
demonstrative-noun (DemN)	Eurasia, excluding MSEA and northwestern Europe	Dryer 2013e
numeral-noun (NumN)	Eurasia, excluding parts of MSEA	Dryer 2013g
relative clause-noun (RelN)	Eurasia, excluding parts of MSEA, parts of South Asia, Europe, and the Near East	Dryer 2013h
degree word-adjective (DAdj)	Eurasia, excluding parts of MSEA	Dryer 2013d
non-initial interrogative	Eurasia, excluding Europe	Dryer 2013k
morphological case marking	Eurasia, excluding western Europe and MSEA	Iggesen 2013
suffixing in inflectional morphology	Eurasia, excluding MSEA	Dryer 2013m
m-T-pronouns present	large parts of Eurasia, excluding MSEA and its adjacent areas to the northeast and west	Nichols & Peterson 2013

But Mandarin also differs from most of Sinitic in structure. Southern Sinitic exhibits stronger affinities to Southeast Asian languages than does Mandarin, which has been more strongly influenced by languages in NEA. There is a debate as to whether the special structure of Mandarin can be explained by “Altaicization”, i.e. influence from Turkic, Khitano-Mongolic, and Tungusic (Hashimoto 1986), or reduction due to non-native acquisition of speakers of languages in nowadays northern China (McWhorter 2007: 104–137). But in any case, this can be labeled an areal feature that separates Mandarin from the rest of Sinitic. Following an extensive discussion, de Sousa (2015: 429) concludes the following:

Some studies on the MSEA linguistic area leave out the languages in China. This is unwise, as the centres of diversity for the Kra-Dai and Hmong-Mien families are still in Southern China, and the Southern Sinitic languages also have many MSEA linguistic traits. Studies of the MSEA linguistic area would benefit immensely if the Southern Sinitic languages, the Far-Southern Sinitic languages in particular, are included in the MSEA linguistic area.

Within the human genome, too, there is a marked difference between Northern and Southern Han populations, the dividing line of which roughly coincides with the Yangtze river (e.g., Zhao Yong-Bin et al. 2014 and references therein). As is well-known, there is also a stereotypical division into North and South as perceived by the Chinese themselves that at least in part has a basis in actual facts such as the predominant cultivation of wheat and rice, respectively (e.g., Eberhard 1965: 601f.). My approach thus stands opposed to Heggarty & Renfrew (2014a: 870), who classify the linguistic landscape of *East Asia* around a “Chinese core” into a northern, a Sinitic, and a southern zone. Of course, all Sinitic languages share certain inherited properties. Perhaps, Sinitic and especially Mandarin may thus be better conceptualized as a transitional zone between MSEA and NEA (Dryer 2003: 48ff.; Comrie 2008). However, in stark contrast to Northeast Asia, Mainland Southeast Asia (MSEA) generally has the following word order features: SVO (SV & VO), AdpN, NGen, NAdj, NDem, NNum, NRel, AdjD (de Sousa 2015: 366). Languages in MSEA usually lack inflectional morphology and have no sign of m-T-pronouns. Of the features listed in Table 3.1, MSEA only shares the non-initial interrogatives. However, for this southeastern neighbor a much longer list of distinguishing linguistic features, such as the lack of a voiced [g] or the existence of complex tone systems, has been summarized by Enfield & Comrie (2015: 7f.). At least for some of them there is no clear-cut boundary to neighboring areas. For instance, Mandarin, Manchu, and Japanese share a similar syllable structure with only very few possible final consonants. In Manchu the only exceptions are ideophones, which is yet another feature that is not unique to MSEA but shared with many languages in NEA as well.

### 3.4 Northeast Asia

In terms of language diversity and phylogenetic diversity, MSEA and NEA show strikingly different patterns as well (Table 3.2). Southeast Asia is home to only five language

### 3 Areal typology and Northeast Asia

families, but in its broadest definition encompasses almost 600 languages. During the preparation of this study it became increasingly clear that an exact number of languages cannot possibly be given for NEA. There is a constant fluctuation of languages spoken by tourists, exchange students, foreign workers, etc. But even if one leaves aside this problem, it is by no means clear at what point a dialect should be counted a language or at what point a language should be considered extinct. For instance, the northern Tungusic languages Evenki, Even, Negidal, Oroqen, and Solon, each of which has strong internal dialectal variation, as well as the extinct language Arman form a complex net of dialect continua. If one agrees with the traditional point of view and considers Arman a dialect of Even, the language as such never went extinct (cf. Doerfer & Knüppel 2013). Evenki alone has about 50 different dialects and experts disagree on whether Oroqen dialects should be included in the list or not (e.g., Whaley & Li 2000; Janhunen 2012d: 7). Given the rapid shift of speakers of both Evenki and Oroqen to Russian and Chinese, respectively, it is often only the older generation that can speak the languages. In some cases no fluent speaker is left, but some relics of the language nevertheless remain in the form of individual expressions or passive speakers. Clear-cut distinctions in these cases are neither feasible nor desirable (cf. Langacker 2008: 13). Leaving aside this fluctuation, most of the dialects, and clearly extinct languages, NEA may be estimated to be home to between 120 and 150 languages. However, NEA shows much more diversity in the number of language families than does MSEA.

Table 3.2: Comparison of language and phylogenetic diversity in MSEA (Enfield & Comrie 2015: 6) and NEA, excluding historically attested languages (this study)

	Core MSEA	Greater MSEA	NEA
language families	5	5	14
languages	280	583	ca. 120-150

Of course, (Greater) Mainland Southeast Asia actually encompasses more than five language families if one includes all small language families (or “isolates”) such as (extinct) Kenaboi, Shom Peng (perhaps Austroasiatic), (extinct) Great Andamanese, or Ongan (Jarawa-Onge) (e.g., Hammarström et al. 2016). The phylogenetic diversity of NEA is also much higher than that of the entire landmass to the west. Excluding extinct languages such as Etruscan and the relatively recent migrations from other parts of the world, there are only representatives of five language families in Europe today, namely Indo-European, Uralic, Basque, Afroasiatic (Maltese), and Turkic. The Caucasus alone adds three more families, but even so, NEA still exhibits much more phylogenetic diversity. Anderson (2010: 137) goes so far as to call the eastern part of NEA, where representatives of 12 of the 14 language families are spoken, a language hot spot with a “high level of unique phylogenetic linguistic diversity endemic to the region”. Of course, if a macro-family such as Transeurasian (Robbeets 2015), allegedly including five different language families, was to be proven, the phylogenetic diversity of NEA would be lower but still

higher than in MSEA, not to mention that there are attempts to lump together language families in MSEA as well (e.g., Sagart 2016). But linguistic diversity in NEA and around the globe is in retreat as many speakers are shifting to larger languages. Not only the number of languages is fading (decrease in language diversity), but whole families such as Ainuic, Amuric, Tungusic, Yeniseic, Yukaghiric, and perhaps Chukotko-Kamchatkan as well as Samoyedic will probably not survive this or the next century (decrease in phylogenetic diversity). Eskaleut, which will persist in other parts of the world, could disappear from NEA as well. In other words, NEA could be the home of languages from only six families in future times, although globalization will bring many more languages from around the world into this area as well.

A good overview of some areal traits found throughout Northeast Asia and adjacent areas has recently been given by Nichols (2010: 366):

Interior Asia has been a center of language spread at least since the Neolithic. The linguistic evidence points to strong and long-term areality in the epicenter of spread, with innovations made in the center eventually showing up farther away. To judge from its distribution, the *m-T* pronoun type may have spread early and then developed its strong structural parallelism in later innovations in the center; case–number coexponence is found at the far peripheries of the area (besides Uralic and Indo-European it also occurs in Chukchi and West Greenlandic), but for at least the last few millennia the classic agglutinating type (with monoexponential and transparently segmentable suffixes) has predominated in the epicenter. Phonemic front rounded vowels may have spread from the epicenter more recently. The consistently head-final morphosyntax of Uralic, core Altaic, Japanese, etc. is more generally widespread in Eurasia and not specific to this northeastern area.

In fact, perhaps one of the strongest features of NEA are the front rounded vowels *ü* and *ö*. A previous study by Maddieson (2013) has shown that these are, by and large, restricted to Eurasia, but it seems that this is a relatively late expansion out of NEA where the highest concentration of languages with these vowels can be found (Table 3.3). In many cases, the available descriptions are not extremely specific about the exact nature of the vowels, i.e. whether they are exactly [y] and [ø] or slightly different sounds.

Table 3.3: Front rounded vowels in Northeast Asia in comparison with Maddieson’s (2013) global sample; see §6.4 and the Appendix for the data

Type	NEA	Global
not present (or unknown)	47	525
high and mid	24	23
high only	8	8
mid only	4	6
Total	83	562

### 3 Areal typology and Northeast Asia

The comparison of the two different samples, global and Northeast Asian, is quite revealing. While altogether 36 out of 83 languages in NEA have at least one kind of front rounded vowel (about 43%), Maddieson (2013) found only 37 out of a sample of 562 languages (about 7%). There are almost no languages of this type along the Pacific Rim, i.e. in Pacific NEA. In fact, excluding the far Northeast (Eskaleut and Chukotko-Kamchatkan) as well as Japan (Japonic and Ainuic) results in an even larger number of 60% (36 out of 60 languages). In NEA all languages with front rounded vowels are from seven language families, namely Koreanic, Khitano-Mongolic, Trans-Himalayan (especially Sinitic), Tungusic, Turkic, Uralic, and Yukaghiric. They were historically lost in many Mongolic and especially Tungusic languages, in the latter case possibly because of contact with languages along the Pacific Rim such as Amuric. Maddieson (2013) mentions only a few languages outside of NEA with front rounded vowels. Of these, four in the Americas, three in the Pacific region and one in Africa are of no concern for us here. But there are several languages in Eurasia, more exactly, six to the adjacent south and thirteen to the west of NEA that also share the phenomenon. Interestingly, the languages in the west include many that have an origin further to the east or within NEA (Hungarian, Finnish, Mari, Turkish, Azeri, Bashkir, Chuvash). Table 3.4 summarizes whether front rounded vowels can be reconstructed to the fourteen proto-languages of languages that are today located in NEA. There will be no comment on the accuracy of the reconstructions and on the details of later developments here, which goes beyond the possibilities of this study. But it may be noted that Vovin's (1993) reconstruction of Proto-Ainuic in this case is highly doubtful.

Table 3.4: Reconstructed front rounded vowels (FRV) for languages spoken in NEA

Language	High	Mid	Source	Comment
Proto-Amuric	-	-	Fortescue 2011: 4	
Proto-CK	-	-	Fortescue 2005: 6	
Proto-Eskaleut	-	-	Fortescue 1998: 125	
Proto-IE	-	-	Fortson 2010: 66	
Proto-Japonic	-	-	Pellard 2008: 136	
Old Korean	-	-	Whitman 2012: 28	Korean has both
Old Chinese	-	-	Baxter & Sagart 2014a: 195	Mandarin has high FRV
Ket	-	-	Georg 2007: 61	
Proto-Ainuic	?+	?+	Vovin 1993: 42ff.	?verbal stems only
Proto-Uralic	+	-	Sammallahti 1988: 481	
Proto-Samoyedic	+	+	Janhunen 1998: 463	
Proto-Mongolic	+	+	Janhunen 2003d: 4	later both mostly lost
Khitan	+	-	Miyake 2017: 493	
Proto-Tungusic	+	+	Doerfer 1978a; 1978b	later both mostly lost
Proto-Turkic	+	+	Róna-Tas 1998: 70	
Proto-Yukaghiric	+	+	Nikolaeva 2006: 57	high FRV later lost

I currently lack exact reconstructions for Koreanic, Trans-Himalayan and Yeniseic, which is why Old Korean, Old Chinese and Ket have been listed instead. The presence of front rounded vowels in Yukaghiric corroborates the hypothesis that this language family historically derives from a location further to the south (§2.14). Similarly, Uralic likely derives from a location close to or perhaps even in NEA (§2.12). Front rounded vowels in this part of the world seem to be a “Ural-Altai” phenomenon (including Yukaghiric but excluding Japonic and Koreanic). However, the origin of the similarity does not necessarily lie in a common origin but may well be the result of prehistoric language contact in southern NEA. It seems that the historical center of the phenomenon clustered around Lake Baikal. Perhaps, its emergence is connected to the phenomenon of vowel assimilation, i.e. vowel harmony (e.g., Maddieson 2013: Chapter Text). The history of German shows that vowel assimilation (in this case umlaut) can most likely be responsible for the emergence of front rounded vowels. The list of proto-languages with front rounded vowels roughly corresponds to the list of proto-languages with KIN-interrogatives. Exceptions include Tungusic (without KIN-interrogative), and Eskaleut (without front rounded vowels). The status of both the interrogative (*hunna* ‘who’) as well as the vowels in Proto-Ainuic is questionable.

A well-known concept of areal linguistics is that of *spread* versus *residual zones* (e.g., Nichols 1992: 13–24; 1997; 2015; Dahl 2001: 1460f.; Bisang 2010: 431f.). Large parts of NEA, especially in the steppes towards the west and along the Lena qualify as spread zones (Nichols 1992: 13–24). In fact, the Eurasian steppe was her prime example. Spread zones are areas with low phylogenetic diversity, low structural diversity, and also low language diversity per language family. There is also no accumulation of diversity over time. In spread zones there is rapid expansion of languages over vast areas that subsequently serve as *lingua francas* for and often replace languages previously spoken in that area.

Each language or dialect group spreading westward on the steppe probably took the form of a classic dialect-geographical area, with a center of innovation (in its eastern range, at least initially) and archaisms on the periphery. Certainly there were centers of political, economic, and cultural influence (Nichols 1992: 16)

One prime example of language spread is the expansion of Sinitic from around the Yellow River southwards towards MSEA, an event influenced by state building, complex social structures, and warfare. Beginning in the 18th century, Mandarin, again starting from about the same area, expanded towards the regions around core China, i.e. Manchuria, Inner Mongolia, Xinjiang, Tibet, Qinghai, and the Southwest. Mandarin is not only used as a main language of communication in all of China and is rapidly replacing many minority languages, but is currently also influencing or even replacing several Sinitic varieties in the South that are the result of the earlier spread. The history of the southern parts of NEA over thousands of years is strongly based on the emergence and spread of multicultural and multilingual confederations ranging from even before the ancient Xiongnu (ca. 3rd century BCE to 4th century CE) to the Manchus from the 17th century onward. The moving factor behind the spread of languages and language families can often be found in cultural or technological innovations, the domestication of different plants and animals, etc. In the case of Indo-European (except Anatolian), for instance,

### 3 Areal typology and Northeast Asia

this possibly was the use of the wheel and wagon (Anthony & Ringe 2015). NEA has seen a variety of spreads of languages or language families over large distances, but to my knowledge, in most cases they have not been clearly linked with such innovations yet. We do not know, for example, which language group was connected with the original domestication of the reindeer in NEA several thousand years ago, which happened independently of the domestication in northern Europe (Røed 2008). But we know that the expansion of some northern Tungusic languages, some Samoyedic languages, Yukaghiric, and Chukchi were likely connected with this innovation (e.g., Janhunen 1996: 61ff.; Helimski 1998: 480; Anderson 2006a,e). Further to the south, the domestication of the horse about 7000 years ago was crucial for the steppe cultures, connected with several language families including Indo-European, Turkic, or Mongolic (Anthony 2007). The yak played a comparable role for the high altitude regions in the southern periphery of NEA around the Tibetan highland, but reaching as far north as the Altai (Wiener 2013). The domestication of the dog may have a relatively long history as compared to that of the other animals mentioned. A recent study found evidence “that sled dogs could have been used in Siberia around 15,000 years ago” (Pitulko & Kasparov 2017: 491). In NEA dog sleds were used, for instance, by the Nivkh and some surrounding Tungusic populations, but also by Samoyeds, Yukaghirs etc. However, the spread of languages is not necessarily based on the spread of its speech community by means of growth and migration. Another important mechanism of language spread is *language shift*, i.e. the shift of a given speech community from one language to another (e.g., Nichols 1997: 372; Janhunen 2007b: 74). Most cases are a combination of these factors.

Spread zones are opposed to *residual* or *accretion zones* (Nichols 1997: 369f.), which Nichols (1992: 13–15) illustrated with the help of the Caucasus. These are areas that have greater phylogenetic, language, and structural diversity. Language families tend to be older (i.e., the age of the respective proto-language lies further in the past) and there are fewer movements of peoples and languages than in spread zones. “As in mountain areas, innovations arise in the periphery (in the lowlands) and archaisms are found in the interior (in the highlands).” (Nichols 1992: 14) Residual zones are areas of retreat rather than spread, usually do not show a single *lingua franca* over the entire area, and have an increase of diversity. There are several possible residual zones in, or rather around, NEA, including most of Pacific NEA (e.g., Ryūkyūan Islands, Hokkaidō, Sakhalin, Kuril Islands, Kamchatka, Aleut Islands), the lower Amur, and many mountain ranges and high altitude regions (e.g., Yunnan, Amdo, the Tibetan Plateau, the Himalayas, the Pamir, the Altai). It should be borne in mind that the features of spread and residual zones mentioned above do not all apply in every case but represent valid tendencies.

Both Anderson (2006d) and Pakendorf (2010) grant Northern Tungusic (more precisely Ewenic) languages a special position for the Siberian area

The features of the Siberian linguistic macro-area cluster around those of the Northern Tungusic languages and this is not by accident. Indeed, the highly mobile Evenki (and to a lesser degree its sister language, Even) both have the local bilingualism relationships and widespread distribution necessary to make them likely vectors of diffusion for at least some of these features (Anderson 2006d: 294)

Expanding on this proposal, one might argue that it is not only the Ewenic branch, but all of Tungusic that used to have a rather special position for NEA. Today, Tungusic languages are mostly endangered, moribund or already extinct (Janhunen 2005; Tsumagari et al. 2007), but one should not underestimate their historical influence over all of NEA. Most Tungusic languages are still located in Manchuria where they had a certain amount of impact on Mongolic, Amuric, Ainuic, and Koreanic. While Evenki and Even expanded into northern NEA and reached places as far apart as Kamchatka and the Taimyr Peninsula, Jurchen and Manchu played an important role for the southern half. The Jurchen established the Jin-dynasty (1115-1234) in northern China and the Manchus had an even more pronounced influence during their Qing-dynasty (1636-1912) that at its height not only included all of modern China, but also what is now the Russian Primorye region as well as Mongolia. Most importantly, Manchu played the role of an ad- and substrate language of Mandarin in Peking, which later was the basis of Standard Mandarin. Of course, there are other language families such as Khitano-Mongolic or Turkic that had an even stronger impact in large parts of NEA.

### 3.5 Subareas in Northeast Asia

The following gives a brief overview of areas of linguistic convergence and contact found in NEA. The areas may overlap strongly with each other, which is not indicated in every case. The discussion limits itself to those areas that seem to be most important for this study.

The territory of NEA as defined here is covered by six different countries: China, Russia, Mongolia, North Korea, South Korea, and Japan. Each of these countries has a **national language** that increasingly influences or even replaces all other languages and dialects within that country. In the case of North and South Korea these are almost identical. There are, therefore, five superstrate languages that may be seen as defining special kinds of linguistic areas. The expansion of the national languages proceeds at the expense of languages and dialects alike. But even if a given language is strong enough to resist the complete loss (e.g., Amdo Tibetan, Buryat, Tuvan, Uyghur, or Yakut), it is usually heavily influenced by the national language. Because the superstrate is the same in the entire country, there is thus a general tendency for all languages to become more similar to each other.

Apart from the areas of Mandarin and Russian influence, **Siberia** is doubtless the largest subarea of NEA. There have been many studies on this northern half of NEA (e.g., Fortescue 1998; Anderson 2004; 2006d; Skribnik 2004; Vajda 2009b; Comrie 2013), but its status as a linguistic area has not been finally clarified. The best summary of areal features found throughout Siberia has been given by Anderson (2006d): vowel harmony, a high back unrounded vowel [u], four nasals (*m, n, ñ, ŋ*), an initial velar nasal (*ŋ-*), SOV word order, morphologically marked reciprocal and desiderative, converbs, case-marked clausal subordination, many cases (especially a prolativ), suffixing morphology, a distinction of dative and allative (but see Pakendorf 2010: 715). Quite problematically, many of the features such as SOV word order or suffixing morphology are not specific to Siberia

### 3 Areal typology and Northeast Asia

(§3.2). The velar nasal indeed shows a very interesting areal pattern, but fails to define a Siberian area as well (Anderson 2013). The Caucasus as well as parts of Europe and South Asia tend towards the absence of the velar nasal altogether. An extremely large area covering most of central Eurasia has velar nasals but not in initial position. Crucially, not only the northern parts of NEA, but also MSEA have the velar nasal in initial position. Nevertheless, NEA has a sharp boundary to the Americas where a majority of languages lack a velar nasal altogether. Interestingly, Mandarin and Manchu, both located in the southern half of NEA, historically lost the initial velar nasal, perhaps due to contact with Mongolic or Turkic. Despite the addition of other possible features such as the special use of speech act verbs (Matić & Pakendorf 2013), Siberia clearly does not qualify as a linguistic area comparable with Mainland Southeast Asia (Comrie 2013). Furthermore, a treatment of Siberia without the inclusion of at least parts of Manchuria and Mongolia is necessarily incomplete.

Several scholars, notably Comrie (1981: 261–266), Anderson (2003), and Georg (2008), have pointed out the special position of **Yeniseic** languages in Siberia. Comrie (2003: 8) summarized the typological differences as follows.

Ket alone has phonemic tone, and Ket alone has a consistent gender/class system, distinguishing masculine, feminine, and neuter nouns, with assignment to masculine and feminine genders using semantic features that go well beyond a mere male/female distinction. While most neighboring languages have relatively simple, agglutinating morphological structure, Ket has a substantially different system, making use, for instance, of internal flexion and discontinuous roots; and while the neighboring languages are at least primarily suffixing, Ket makes widespread use of both suffixes and prefixes.

However, some languages, notably Middle Mongol, Khitan, Manchu, and perhaps even the mysterious language of the Rouran (Vovin 2004), had a limited gender or sexus system. Prefixes can, furthermore, also be found in Ainuic, Yukaghiric, or Chukotko-Kamchatkan. Tones are also present in Japonic, Koreanic, and Trans-Himalayan.

**Altaic** is perhaps one of the most disputed proposals for a language family worldwide. Still today, there is complete disagreement over the validity of the family and no conclusion is in sight. Robbeets (2015) restricts the name *Altaic* to Turkic, Khitano-Mongolic, and Tungusic and uses the term *Transeurasian* as a cover term for Altaic, Japonic, and Koreanic. However, given that Altaic or Transeurasian fails to be accepted by a majority of scholars, it must be considered an *unproven* (but still interesting) hypothesis. This does not mean, however, that it is not possible that some of the languages, say, Khitano-Mongolic and Tungusic are ultimately related to each other. Janhunen (1996) proposed the name *Khinganic* for the hypothetical language family that unites Tungusic and Khitano-Mongolic. It should also be pointed out that there are more possibilities than genetic relatedness on the one hand and mutual contact on the other hand (cf. Doerfer 1985). One of several imaginable scenarios is that at least one of the proto-languages involved really was a mixed language and thus had no clear-cut affiliation in the first place. Any theory would need to explain the fact that Turkic, Tungusic, and Mongolic share

similarities in the pronominal system that are both too similar to be due to chance and at the same time too similar to be of common origin, especially given the absence of a common inherited vocabulary (Janhunen 2013: 221). An explanation that seems to be gaining acceptance sees the observable lexical similarities as borrowing, especially from Turkic to Mongolic and from Mongolic to Tungusic (e.g., Doerfer 1985; Schönig 2003). Therefore, one of the most important tasks still is the identification of layers of loanwords in all five language families (e.g., Khabtagaeva 2017; p.c. 2018). A bilateral relation of two of the so-called Transeurasian languages, namely Koreanic and Japonic, is being advocated independently of their relation to Turkic, Khitano-Mongolic, or Tungusic (Whitman 2012), but awaits further discussion. Altaic in some parts even today is rather premature in the sense that the internal reconstruction of the individual language families should still take priority. For instance, it does not make too much sense to compare Proto-Mongolic with Proto-Tungusic before the evidence by Khitan (e.g., Janhunen 2012a), other Para-Mongolic languages (e.g., Shimunek 2017), the Hüis Tolgoi inscription (e.g., Vovin 2017), and less well-known Tungusic languages is taken into account. The necessary first step for the Khitano-Mongolic side must be the continuing decipherment of Khitan, followed by, if possible, an improved reconstruction of Proto-Khitano-Mongolic. On the Tungusic side, evidence from several languages such as Alchuka, Bala, or Kyakala (e.g., Mu Yejun 1985; 1986; 1987, Hölzl 2017b; 2018b) as well as the dialects of Oroqen (Whaley & Li 2000) and Manchu (e.g., Hölzl 2018a) keep being neglected in most studies.

The term *Ural-Altaic* was originally a proposal for a language family which has long since been abandoned. Janhunen (2007b: 78) revived the term in an areal typological sense as

a complex of several language families covering the entire trans-Eurasian belt from Finland and Lapland in northern Europe to Korea and Japan in the Far East. Other regions where Ural-Altaic languages are spoken, or have until recently been spoken, include Pannonia, Anatolia, Western and Eastern Turkestan, Mongolia, Manchuria, much of Russia, and most of Siberia. The language families conventionally ‘classified’ as Ural-Altaic are: Uralic, Turkic, Mongolic, Tungusic, Koreanic, and Japonic.

This is as problematic as the use of *Altaic* as a typological label (Janhunen 2007b; Janhunen 2007a): “areal typology would study the geographical distribution of such features, rather than the characteristics of individual areas.” (Dahl 2001: 1456) In fact, several of the few features mentioned by Janhunen are not characteristic of “Ural-Altaic” languages only, but can be found in the larger Eurasian area (§3.2). A label such as “Altaic” or “Ural-Altaic” may not only lead to misinterpretations concerning genetic connections, but also suggests a certain typological homogeneity, which—as Janhunen is clearly aware—is not always the case. In many cases such as the pronominal similarities it is more plausible to add Yukaghiric to Ural-Altaic rather than Japonic or Koreanic.

To my knowledge, Janhunen (1996; 1997) offers the only explicit treatment of languages in **Manchuria**. But whatever the exact delimitation of Manchuria, it certainly does not qualify as a linguistic area in any sense because it has not been defined on

linguistic grounds at all. There are, however, several areas of strong convergence and bilingualism within Manchuria such as one around the Mongolic language Dagur that includes the Tungusic languages Solon, Oroqen, Manchu, and Mongolian dialects such as Khorchin. Gusev (2015b: 72) argues for a linguistic area “which includes the dialects of Negidal, Amur Tungusic, Nivkh and Ainu, and in some respects may be a part of a larger area, that could embrace other varieties, such as Evenki and Even, Hokkaidō Ainu, Japanese and the languages of Kamchatka”, but fails to mention any defining features for this larger area altogether (but see also Yamada 2010). For interference between different Tungusic languages in Manchuria see §5.10.1 and §6.3.

Usually considered one of the best examples of a linguistic subarea in NEA, the **Amdo Sprachbund** can be found in northwestern China (e.g., Dwyer 1995; Slater 2003a; Zhong Jinwen 2007). There are many different designations for the area (Janhunen 2007a), but the name *Amdo Sprachbund* has been adopted by several recent publications (e.g., Janhunen 2012c; Simon 2015; Sandman & Simon 2016). It is difficult to establish a clear boundary of the area, but it roughly encompasses eastern Qinghai, parts of northern Sichuan, and most of Gansu. The best overview of the area has been given by Janhunen (2007a; 2012c), according to whom the area is the result of a very unique interaction of Turkic, Mongolic, Tibetan, and Sinitic languages. However, historically the Tangut language, usually classified as Qiangic, as well as the probably Para-Mongolic language Tuyuhun were at some point also spoken in the area. Janhunen (2012c: 180ff.) mentions the following defining features of the area: SOV word order, suffixes or enclitics, case marking, verbal tense-aspect categories, converbs, postpositions, indefinite articles, perspective marking (including loss of person marking in Turkic and Mongolic). From this brief list alone, however, this quite clearly does not qualify as a linguistic area at all. Most of these features are not only prevalent in adjacent areas, but are also extremely common worldwide. But it may be noted that, apart from strong interference between individual languages with each other (e.g., Sandman 2012; Simon 2015; Sandman & Simon 2016), there are also several creolized languages such as Gangou, Wutun, Tangwang, and Hezhou Chinese, which indicates strong language contact in the area (§5.9). Before this background it seems even possible to extend the area towards the south to include, for example, the language Daohua, which is a Chinese-Tibetan creole or mixed language spoken in western Sichuan province (e.g., Acuo Yixiweisa 2001; Chen Litong 2017). Nevertheless, the traditional conception of the Amdo Sprachbund is adopted here for pragmatic reasons (see §6.3).

## 4 The typology of questions

Before Chapter 5 investigates the grammar of questions in the individual language families of NEA, this chapter describes the most important parameters of the typology.

### 4.1 Introduction to the typology of questions

There is a certain amount of confusion surrounding the terminology employed for what was called the *grammar of questions* in this study. Grammar books usually employ the terms *question* and *interrogative* (nominal or attributive), but quite inconsistently so. In most cases no clear-cut distinction is drawn and the terminology is simply tacitly taken for granted. A few examples should suffice to illustrate the extent of the problem in English-language publications. Schulze (2007: 250), for example, explicitly employs the term *interrogativity* for the cognitive side of the phenomenon and *question* for the linguistic form. A related terminology can be found in Rajasingh (2014: 103): “*Interrogation* is a semantic process of eliciting information by way of *questioning*.” (my emphasis) Huddleston (1994: 411), on the contrary, “explores the relation between *interrogative*, a category of grammatical form, and *question*, a category of meaning.” (my emphasis) Dixon (2012: 376) draws a distinction between different speech acts (e.g., *questions*) and grammatical categories (e.g., *interrogative*). Furthermore, for what is traditionally known as an *interrogative pronoun* he employs the much more fitting term *interrogative word*. In this study, the term *question* refers either to the formal side or is used as a cover term for both the formal and the semantic side taken together. The semantic side of questions will only be named *interrogativity* or *interrogation* if a clear distinction is called for. The so-called *interrogative words*, following Diessel (2003), will simply be called *interrogatives* in order to preserve a connection to the traditional term and to place at the same time an additional emphasis on their similarity to so-called *demonstratives* and on their possible special position in the language.

For all we know, question-response sequences (Enfield et al. 2010) and, more generally, turn-taking in conversation (Stivers et al. 2009) provide a universal enchronic infrastructure that allows a comparison of different languages with each other. Question-response sequences are usually accompanied by non-linguistic cues such as the gazing behavior of the questioner (Rossano et al. 2009) or head movements by the addressee such as a head shake. For practical purposes, this study concentrates on the first part of such sequences exclusively, and must leave aside non-linguistic elements. While this omission will perhaps cause some eyebrow-raising among experts, such information can only be obtained through prolonged fieldwork and thus is only available in sufficient detail for very few languages worldwide (e.g., Levinson 2010).

#### 4 The typology of questions

A full account of the historical development of the typology of questions lies beyond the possibilities of this study. In the following, I will only give a rough sketch with a focus on more recent advances. Apart from some isolated and mostly outdated studies (e.g., Bolinger 1957), the modern typology of questions by and large may be said to have started around 1970 with Ultan (1978), a cross-linguistic study based on a sample of 79 languages (originally published in 1969), Moravcsik's (1971) investigation of polar questions in 85 languages, and Danielsen (1972), based on a sample of about 60 languages. Since then, the field has made enormous advances that cannot be reviewed here in every detail. During the 1970s and 1980s there were relatively few *important* publications with long-lasting effects, such as a collection of papers on questions in seven languages in Chisholm (1984) and the study by Sadock & Zwicky (1985) (written around 1976 and 1977) in the first edition of *Language Typology and Syntactic Description*. The number of works has been steadily increasing at an ever faster pace from the 1990s until today. By now there are several dozen important publications, not including studies on individual languages, the number of which has been growing even more rapidly. But surprisingly, the only investigation that may be said to represent something like a standard typology is Siemund (2001), which is a mere 19 pages long and by now over 15 years old. A somewhat updated account by König & Siemund (2007) can be found in the second edition of *Language Typology and Syntactic Description*. Perhaps the best general introduction to the typology of questions to date can be found in volume three of Dixon's (2012: 376–433) *Basic Linguistic Theory*. Table 4.1 gives a non-exhaustive overview of some important typological studies of questions since 1990, excluding investigations of individual languages and generative approaches. Few studies are based on a large sample and almost all are unrepresentative of the languages of the world. Exceptions include Idiatov's (2007) lengthy investigation of 1850 languages and especially a series of high-quality investigations with a sample of about 900 languages by Dryer (2013l,k,j). Most studies only focus on specific details but do not cover the entire scope of the grammar of questions.

There are many possible classifications of questions. For instance, Sanitt (2007: 439) draws a distinction between *empirical* ("questions whose presuppositions are undoubted or taken as axiomatic") and *theoretical* questions ("all questions which are not empirical"). Sanitt (2011: 559) furthermore introduces a distinction between *closed* questions that "have definitive answers" (such as a riddle, see §4.4) and *open-ended* questions that "lead to other questions". These distinctions may be useful for the philosophy of science (e.g., Meyer 1980), but to the best of my knowledge they are not relevant for a cross-linguistic investigation.

The typology proposed in this study differs from many previous typological accounts that usually first drew a distinction between different question types, especially polar and content questions. A similar focus on polar and content questions can also be found in most grammar books and specialized descriptions of questions in individual languages. In contrast, the present study makes a *primary distinction* between (1) question marking, (2) interrogatives, and (3) other functional domains such as focus that can combine with question marking or interrogatives. Only a secondary distinction is made within the domain of question marking in different question types (Hölzl 2016b). Of course, this is

#### 4.1 Introduction to the typology of questions

Table 4.1: Important typological studies of questions since 1990

Authors	Focus	Languages	Number
Muysken & Smith 1990	interrogatives	pidgins and creoles	ca. 25
Heine et al. 1991	interrogatives, hierarchy	global, unrepresentative	14
Lindström 1995	interrogatives	global, unrepresentative	24
Mushin 1995	interrogatives	Australian languages	26
Nau 1999	interrogative paradigms	European, Australian languages	ca. 20?
Huang et al. 1999	general	languages of Taiwan	7
Siemund 2001	general	global	ca. 50?
Bencini 2003	question marking, grammaticalization	global, unrepresentative	ca. 25?
Diessel 2003	interrogatives	global, unrepresentative	100
Bhat 2004	interrogatives	global, unrepresentative	ca. 80?
Idiatov & van der Auwera 2004	interrogatives	global, representative	ca. 350
Hackstein 2004	rhetorical question, grammaticalization	Indoeuropean languages	?
Cysouw 2005	interrogatives	global, unrepresentative	?
Cysouw 2007	interrogatives, grammaticalization	Pichis Ashéninca (Arawakan), global, unrepresentative	?
Idiatov 2007	interrogatives (who, what)	global, representative	1850
König & Siemund 2007	general	global, unrepresentative	?
Schulze 2007	general	global, unrepresentative	?
Lichtenberk 2007	interrogatives	Oceanic languages	ca. 25
Hagège 2008	interrogative verbs	global, representative?	28 (217?)
Mauri 2008	question marking, alternative questions	global, representative	ca. 60?
Mackenzie 2009	interrogatives, hierarchy	global, unrepresentative	50
Rialland 2009	intonation	African languages	145
Stivers et al. 2009	turn-taking	global, unrepresentative	10
<i>Journal of Pragmatics</i> 42 (e.g., Levinson 2010)	general	global, unrepresentative	10
Axelsson 2011	tag questions	global, unrepresentative	ca. 12?
Miestamo 2011	polar question marking	Uralic languages	ca. 30 (200?)
Dixon 2012	general	global	ca. 30?
Hengeveld et al. 2012	interrogatives	languages of Brazil	24
Dryer 2013l [2005]	polar question marking	global, representative	884
Dryer 2013k [2005]	interrogatives, position	global, representative	902
Dryer 2013j [2005]	polar question marking	global, representative	955
Luo Tianhua 2013	general	languages of China	138
Hackstein 2013	polar question marking, grammaticalization	Indoeuropean languages	ca. 10
Dingemanse et al. 2013	the word <i>huh</i>	global, unrepresentative	31
Haspelmath & the APiCS Consortium 2013b	interrogatives, position	pidgins and creoles	76
Haspelmath & the APiCS Consortium 2013a	question marking	pidgins and creoles	76
Köhler 2013	question marking	Ometo languages	6
O'Connor 2014	interrogatives, position	south American Indian languages	44
Hölzl 2015e	question marking	global, unrepresentative	50
Bowern et al. 2016	general	hunter-gatherer languages	194
Hölzl 2016b	question marking	global, unrepresentative	ca. 20
Köhler 2016	general	African languages	119?

#### 4 *The typology of questions*

not an altogether new endeavor. Similar ideas have already been formulated, for example, by Bhat (2004: 248-249) for content questions.

The purpose of using these pronouns [interrogatives ~ indefinites] in such sentences is merely to indicate that the speaker lacks knowledge regarding a particular constituent. There are two other meanings that need to be expressed in constituent questions [CQ], namely (i) a request for information (interrogation) and (ii) restriction of that request to a particular constituent (namely the indefinite pronoun); these meanings are generally expressed, in these languages, with the help of additional devices; for example, devices like the use of question particles or question intonation are used for denoting interrogation, whereas devices like the use of focus particles or focus constructions are used for denoting that the interrogation is restricted to a particular constituent. (my square brackets)

Bhat's approach was the impetus for a primary separation of question marking from interrogatives and focus in this study. However, while Bhat concentrated exclusively on content questions, this study includes further question types such as polar, alternative, and focus questions.

This typology excludes echo, rhetorical, and indirect questions. For the most part, research commonly known as "wh-movement" or "wh-fronting" (e.g., Cable 2007) will not play a dominant role within this study either. First of all, very few languages in NEA exhibit this phenomenon that can by no means be said to be a universal property of language. Second, it is, strictly speaking, neither part of the domains of question marking, nor of interrogatives, but belongs to the domain of focus marking. This study for the most part also excludes the grammatical category of indefinites that are usually derived from interrogatives and have been discussed in detail elsewhere (see Van Alsenoy & van der Auwera 2015 for a list of references).

### 4.2 Question marking

Typological variation within the domain of question marking includes (1) different formal marking strategies, (2) different semantic scopes of question markers over question types, (3) interaction of question marking with other functional domains, and (4) the overall number of question markers. But before I investigate these four aspects one after another, let me introduce the major question categories that can take question marking.

The *major question categories* are polar questions (PQ) (also called yes-no questions), content questions (CQ) (also known as constituent or wh-questions), alternative questions (AQ) (or disjunctive questions), and perhaps focus questions (FQ). The designations are somewhat problematic but nevertheless will be adopted here because they are the most common and conventional labels. Consider the following examples corresponding to the declarative sentence *I want coffee* (cf. Hölzl 2016b: 18):

- (1) English
- a. Do you want coffee?
  - b. Do you want coffee?
  - c. Do you want coffee or tea?
  - d. What do you want (to drink)?

Generally, a different question category can be postulated if it has a specialized marking in at least one but preferably more languages. English alone, for example, fails to differentiate polar and focus questions in terms of question marking. Examples (1a) and (1b) exhibit the same question marking but differ with respect to the marking of focus. Focus is understood here in a very broad sense of a “restriction of that request to a particular constituent” (Bhat 2004: 246) and usually has a contrastive function. Alternative questions, though exhibiting the same marking, in addition contain a disjunction preceding the second alternative, which is not the case in polar and focus questions. Now consider the following examples from the Na-Dene language Slavey.

- (2) Slavey (Hare, Na-Dene)
- a. *sú duká ʔanehʔiʔ*  
Q like.this 2SG.do  
‘Do you do it this way?’ (Rice 1989: 1123)
  - b. *duká nɪ ʔanehʔiʔ*  
like.this Q 2SG.do  
‘Is this the way that you do it?’ (Rice 1989: 1133)
  - c. *legóhlini gotsʔé ʔawodee gúshɪ kóé ʔawotʔéʔ*  
PN area.to 2SG.OPT.go or home 2SG.OPT.be  
‘Are you going to Norman Wells or staying home?’ (Rice 1989: 1139)
  - d. Slavey (Mountain, Na-Dene)  
*ʔamí ʔayíláʔ*  
who 3.affected.4  
‘Who did it?’ (Rice 1989: 1141)

The examples in (2a) to (2d) are instances of polar, focus, alternative, and content questions, respectively. Unlike English, there is a clear distinction between polar and focus questions. Polar, content, and alternative questions are generally accepted as separate categories, although alternative questions are sometimes subsumed under polar questions (e.g., Siemund 2001). *Focus questions* (Kiefer 1980), on the other hand, are a category that is not usually recognized or included in grammatical descriptions, but nevertheless has some validity as shown in example (2b) above (see also Dixon 2012: 395-396). Miestamo (2011: 2) includes them in his definition of polar questions “encompassing all

#### 4 The typology of questions

interrogatives eliciting yes/no replies, regardless of whether they are neutral or biased towards a positive or negative answer, or whether they have [a] broad sentence focus or a more narrow focus on a particular constituent.” The prime example for the cross-linguistic relevance of focus questions used in Hölzl (2016b) is the Japonic language Yuwan (see §5.6.2), which contains specialized question markers for polar (*-mi*), focus (*-ui*), and content questions (*-u*). Dixon (2012: 396) mentions yet another example with different polar (*-éé*) and focus question marking (*-áa*), the Cushitic (Afroasiatic) language Tunni. However, he also includes focus questions in the category of polar questions. Miestamo (2011) is certainly correct that in many languages focus questions exhibit affinities with polar questions, but this is not always the case. In the Australian language Bardi (3), for example, there is an affinity between focus and alternative, but not polar questions, which are marked with sentence-initial *nganyji*, which is derived from an interrogative.

- (3) a. Bardi (Nyulnyulan)  
*gooyarr=arda aarli mi-n-Ø-nya-gal?*  
 two=Q fish 2-TR-[PST]-catch-REC.PST  
 ‘Was it *two* fish that you caught?’
- b. *ngay=arda nga.n.k.iid.a broome-ngan, gardi joo=warda?*  
 1MIN=Q go PN-ALL or 2MIN=Q  
 ‘Will I go to Broome today or will it be you?’ (Bowern 2012: 619)

There do indeed seem to be relatively few languages with specialized marking strategies for focus questions, although this might be a distortion due to the fact that the category itself is not widely known in linguistic circles and therefore did not make it into grammatical descriptions.

Following the methodology sketched above, there are some indications for additional types of questions. One such example are *negative polar questions* (NPQ). In most languages, including English, these are expressed in the same way as plain polar questions except for the addition of a negator, e.g. *Don’t you want coffee?* But in Urarina, a language with no clear affiliations to other languages spoken in Peru, there is a specialized NPQ marker *ta* different from the plain PQ marker *=na*. An NPQ in addition requires a negative marker *=ne*.

- (4) a. Urarina  
*hanone mēsahe auna-i=pa?*  
 morning message hear-2=Q  
 ‘Did you hear the message in the morning?’
- b. *ta kure kwituku-i=pe?*  
 Q price know-2=NEG  
 ‘Don’t you know the price?’ (Olawsky 2006: 832, 834)

Note the different syntactic behavior of the plain and the negative polar question markers. This category seems to play no important role for NEA, however, which is why it has not been taken into account in this study (but see §5.6.2 on Shuri).

Minor subtypes of questions include *negative alternative questions* (NAQ, *Do you want coffee or not?*) and *open alternative questions* (OAQ, *Do you want coffee or what?*), but their cross-linguistic relevance remains to be investigated. NAQs are mentioned as a separate category because they play a very important role in the grammaticalization of polar question markers (§4.2.3). The category of open alternative questions has been proposed by Tolskaya & Tolskaya (2008). NAQs and OAQs, it seems, do not fulfill the requirement as an independent question type as they are generally based on the same construction as alternative questions. In this study the two are simply taken as useful labels for special subtypes of alternative questions.

Table 4.2 lists some defining properties of the major question types described above. There may well be additional properties, but these are the most important ones for the purposes of this study. The term *proposition*, which is a very common and useful label, should not be understood in a logical sense, but in terms of *embodied simulation* (see §4.4). Polar questions and focus questions both expect an answer in the positive or in the negative. If they are marked with the same marker, this usually attaches to the verb in polar questions and to the focal element in focus questions. Very often, the same question marker can also be found in alternative questions, where it attaches once to each alternative. Focus questions and content questions share a narrow focus on one constituent. The difference lies in the fact that in content questions this usually is an interrogative, while in focus questions it is usually a fully specified nominal or verbal phrase or other part of the sentence.

Table 4.2: Important properties of individual question types (cf. Dik 1997: 260; Dixon 2012; Hözl 2015d; 2016b).

	<b>Expectancy from hearer</b>	<b>Questioned part of proposition</b>	<b>Level of specificity</b>	<b>Specified alternatives</b>
CQ	specific information	part (or whole)	schematic	none
PQ	confirmation or negation	whole	specific	one
FQ	id.	part	id.	id.
AQ	selection from alternatives	part (or whole)	specific (or schematic)	two or more

However, content questions do not necessarily contain an interrogative as can be seen in Wari', a Chapacuran language spoken in Brazil in which demonstratives fulfill the function of interrogatives (Everett & Kern 2007). The presence of interrogatives as a defining feature of content questions is also problematic otherwise. For instance, there are languages in which an interrogative develops into a polar question marker but is still identical to the interrogative. Content questions are better defined as questions that have a narrow sentence focus on a schematic subpart of the "proposition" and thus inquire about very specific information instead of a confirmation (cf. Dixon 2012). For instance,

#### 4 The typology of questions

*Are you leaving tomorrow?* (FQ) is much more specific than *When are you leaving?* (CQ). The level of specificity has been adopted from Arnheim (1969: 238) and Langacker (2008: 19) and will be further described in §4.4. Open alternative questions are partly schematic.

Polar questions are unique in inquiring about the whole, but there are some alternative questions such as *Is it raining or did someone leave the sprinkler on?* (Sadock & Zwicky 1985: 179) and content questions of the REASON type (*why?*) that are similar in this regard. The answer to a question such as *Why are you leaving?* may either be a whole (*It is going to rain.*) or a partial clause (*Because of the rain.*). In fact, in many languages interrogatives referring to the category of REASON are derived from interrogative verbs. In many instances in NEA these are converb forms of an interrogative verb meaning ‘to do what’, e.g. Manchu *ai-na-me* ‘what-v-CVB.IPFV’. Because verbs usually stand for the whole “proposition”, this is direct evidence that REASON interrogatives may also refer to the whole as well. But most alternative questions and content questions, as well as all focus questions, focus on a subpart of a given “proposition”. This is directly mirrored in overt focus markers that are often found in focus and content questions as well as the fact that the part of alternative questions that is identical in both alternatives (the unfocused part) may fall victim to ellipsis in one of the two alternatives in the majority of languages.

Another difference concerns the number of alternatives that are specified (Dik 1997: 260). Content questions imply many possible alternatives but specify none. All expected alternatives, however, have in common the schematic meaning just mentioned. In other words, an answer to *When are you leaving?* may be *tomorrow, the day after tomorrow, July 4* etc. If the expectancy was wrong, the answer may also be different (e.g., *No, I am not leaving at all.*), but this is something that all question categories have in common (see §4.4). Polar and focus questions imply more than one alternative but have only one specified. Alternative questions by definition have more than one alternative and do not usually imply more (but see example 22 below from Mauwake).

In some cases it is not easy to differentiate between types of question marking and types of questions. Two of the most difficult examples are *tag questions* and so-called *A-not-A questions*, both of which in previous studies have been treated as either a question type or marking strategy of polar questions (see Hölzl 2016b: 20). (5) and (6) are typical examples (based on own knowledge).

- (5) a. English  
You want coffee, **right**?  
b. You want coffee, **don't** you?
- (6) Mandarin  
你要不要喝茶?  
*nǐ yào bu yào hē chá?*  
2SG want NEG want drink tea  
‘Do you want to drink tea?’

The solution proposed by Hölzl (2016b) is to classify A-not-A as elliptic negative alternative questions without overt marking, but juxtaposition of the two alternatives. From this point of view they are neither a marking strategy for polar questions, nor a question type on their own, but are a very special case of alternative questions (see also Clark 1985). The category of tag questions was either not described or are simply non-existent in the majority of languages in NEA and therefore do not play a significant role within this study. But their presence in many languages from other parts of the world, such as Europe, means that they cannot be neglected. Problematically, many investigations of tag questions fail to define them properly. Furthermore, what is usually recognized as a tag question has a plethora of different meanings, which makes them extremely difficult to define from a functional point of view (Mithun 2012: 2166f.). The fact that tag questions can be described in terms of a *question tag* in relation to a so-called *anchor* (Axelsson 2011) makes them unique among all of the question types mentioned above. In fact, the traditional category of tag questions usually consists of two sentences, which is why they do not, actually, qualify as question type at all. Perhaps tag questions thus have to be described in different terms. For reasons that will become clearer in §4.4, tag questions will be treated as a construction type located on a different level of analysis than the other question types. However, for practical purposes their formal properties will be briefly discussed with the other question types in §4.2.1.

The following subsections address marking strategies (§4.2.1), the scope of question marking over different question types (§4.2.2), the interaction of question marking with other functional domains such as focus (§4.2.3), and finally, the overall number of question markers in individual languages (§4.2.4).

### 4.2.1 Marking strategies

Previous accounts of question marking are often restricted to the marking of polar questions (e.g., Miestamo 2011; Dryer 2013l,j). This section takes a broader perspective and investigates question marking strategies in all question types, including, for the sake of completeness, the problematic category of tag questions.

Miestamo (2011), in analogy to his earlier typology of negation, investigates the distinction and symmetry between the marking of polar questions and declarative sentences. Some of the categories in Dryer (2013j) are likewise based on such a comparison. My own typology builds on these approaches and draws a broad distinction between marked and unmarked polar questions and declaratives, which defines the four different types shown in Table 4.3.

Table 4.3: Marking of polar questions as opposed to declaratives (cf. Hölzl 2016b: 21)

	unmarked PQ	marked PQ
marked declarative	Type 1: Sanuma, Sheko	Type 3: Crow, Sabanê
unmarked declarative	Type 2: Yéli Dnye	Type 4: English, Bengali

#### 4 The typology of questions

For Types 1 and 2 consider examples from the Ethiopian language Sheko (7) and Yéli Dnye (8), a language without clear affiliation spoken on Rossel Island.

(7) Sheko (Omotic, Afroasiatic)

- a.  $\eta=m\bar{a}\bar{a}k-\bar{a}-m$ .  
1SG=tell-put-IRR  
'I will tell.'
- b.  $\eta=m\bar{a}\bar{a}k-\bar{a}?$   
1SG=tell-put  
'Shall I tell?' (Hellenthal 2010: 402)

(8) Yéli Dnye

- yi kópu dê d:uu./?*  
that.ANAPH thing 3.IMM.PST.PL make.PL  
'He made it./?' (Levinson 2010: 2743)

Note that Yéli Dnye even lacks a distinction of intonation. Types 1 and 2 are exceedingly rare and are altogether absent from NEA, which is why they will be neglected in this study (see Köhler 2013 for Type 1). Type 4 is by far the most frequent cross-linguistically, followed by Type 3.

The interaction of overt question markers with intonation complicates matters, but this will be ignored for the moment. The South American language Sabanê (9) and Bengali (10) illustrate Types 3 and 4, respectively.

(9) Sabanê (Nambikwaran)

- a. *iney-i-ntal-i*.  
fall-V-PRET.NEUT-DECL  
'(S)he fell.'
- b. *iney-i-ntal-a?*  
fall-V-PRET.NEUT-Q  
'Did (s)he fall?' (Araujo 2004: 205)

(10) Bengali (Indo-Iranian, Indo-European)

- a. *tumi take cenô*.  
2SG 3SG.OBJ know.2.PR.S  
'You know him.'
- b. *tumi ki take cenô*.  
2SG Q 3SG.OBJ know.2.PR.S  
'Do you know him?' (Thompson 2012: 200)

The different morphosyntactic status of the marker is unimportant for this primary distinction.

However, by considering Types 3 and 4 exclusively, there is a variety of different formal types of *polar question marking* (e.g., Siemund 2001; Miestamo 2011; Dryer 2013l,j). Spoken language is one-dimensional. In order to signal certain information such as interrogativity, there are thus limited means available. We may simply modulate the phonation of the speech stream (intonation), change the order of elements in the speech stream (word order), or we may add material (morphosyntax). Among the elements that can be added are affixes, clitics, or free elements such as particles. These may stand either before or after another element (prefixes vs. suffixes, proclitics vs. enclitics, preposed vs. postposed particles). The element with respect to which these question markers can be located may either be the whole sentence or a subpart such as the first constituent or the verb. Affixes are less free in their position than clitics and particles, and usually attach to the verb.

Apart from some exceptions, *intonation* is not normally described in detail for languages in NEA, if it is mentioned at all. Within this study it was impossible to remedy this unfortunate fact, but where possible some rough outlines are sketched (such as falling or rising intonation etc.). Intonation, although not universal, is certainly among the most important ways of marking questions cross-linguistically. However, in the majority of languages, intonation is combined with other markers. In Dryer's (2013j) sample of 955 languages only 173 languages (about 18%) exclusively made use of intonation for polar question marking. In NEA the number is even lower (Chapter 6). Concerning the location and contour of question intonation there are no absolute universals (see Sicoli et al. 2014: 4 and references therein). In fact, generalizations such as final rising intonation in polar questions are not true for individual languages like English (Couper-Kuhlen 2012), let alone from a cross-linguistic perspective. For example, Riialand (2009: 928) describes what she calls the *lax question prosody* found in a relatively large area of central and western Africa, which is generally characterized by "a falling pitch contour, a sentence-final low vowel, vowel lengthening, and a breathy utterance termination produced by the gradual opening of the glottis." Because of the absence of reliable and good information on intonation, this study necessarily focuses on the material aspect of question marking.

Question marking by *word order change* is almost entirely restricted to Western Europe and Indo-European languages (e.g., Hackstein 2013), and is extremely rare from a cross-linguistic perspective (Dryer 2013j). This is a feature of European languages that clearly differentiates them from the rest of Eurasia, including NEA. An example can be found in Finnish (11).

## (11) Finnish (Uralic)

- a. *sä tuu-t*  
 2SG come-2SG  
 'You're coming.'
- b. *tuu-t sä?*  
 come-2SG 2SG  
 'Are you coming?' (Miestamo 2011: 7, 12)

#### 4 The typology of questions

This seems to be a pattern that originates in Germanic languages from where it spread to some Uralic, Romance, and Slavic languages. Following Miestamo (2011: 12), one may assume an original second position enclitic marking questions as well as focus. Such markers normally attach to the fronted verb in polar questions and the loss of this marker quite naturally leaves the fronting of the verb to mark polar questions. We furthermore know that several Indo-European languages had a second position clitic or particle such as =*ne* in Latin (cf. §5.5.2). According to Miestamo this is also what happened in Finnish, which still preserves a second position clitic in other constructions.

- (12) Finnish (Uralic)  
*tule-t=ko sinä?*  
come-2SG=Q 2SG  
'Are you coming?' (Miestamo 2011: 12)

Perhaps, Germanic had a second position clitic comparable to Gothic =*u* (Braune & Heidermanns 2004: 178) that was already lost in other Old Germanic languages. While the loss of the question marker is not actually attested for Germanic, it is for some other European languages such as the Uralic language Pite Saami. Wilbur (2014: 186-187, 244) notes that there used to be a second position question marker =*gu(s)* in Pite Saami that attached to a verb in polar questions and that almost entirely disappeared during the 20th century. Today, polar questions are usually marked by verb-initial word order only. Of course, the development in languages such as Pite Saami may have been influenced by language contact as well.

Many examples of different morphosyntactic markers can be found throughout this section as well as in Chapter 5, which is why no further examples will be given here. A rare strategy is the use of *infixes* such as in Koasati, a Muskogean language spoken in the US (cf. Dixon 2012: 384). In Koasati, questions may be "formed by infixing a glottal stop between the penultimate and ultimate syllables." (Kimball 1991: 301) The Koasati question marker is a true infix -ʔ- that can, but does not necessarily, coincide with a morpheme boundary (Kimball 1991: 302). Similarly rare are *auxiliaries* that are restricted to marking questions (Miestamo 2011: 4). One example stems from the Salish language Halkomelem, which has the auxiliary *lí-*. This should not be confused with auxiliaries encountered in, but not restricted to, questions such as English *to do*, or with interrogative verbs such as *to do what* that are interrogatives and not question markers (e.g., Hagege 2008). According to Hyman & Leben (2000: 593), there are some languages in which questions can be marked with *tones*:

In **Hausa** {Chadic, Afroasiatic}, a L is added after the rightmost lexical H in a yes/no question, fusing with any pre-existing lexical L that may have followed the rightmost H (which is raised somewhat, as are any following L tones whatever their source). As a result, lexical tonal contrasts are neutralized. In statements, [káì] 'head' is tonally distinct from [káí] 'you [masculine]'. But at the end of a yes/no question, they are identical, consisting of an extra-H gliding down to a raised L. In **Nembe** {Ijoid, ?Niger-Congo}, a final lexical L becomes H in statements, and

a final lexical H becomes L in questions. Thus, L-L / LH contrasts such as [diri] ‘book’ / [bùrú] ‘yam’ are neutralized as L-H in statements, but as L-L in questions. A similar case is found in **Isoko** {Atlantic-Congo, Niger-Congo}, where a final L marks positive questions, while a final H marks negative questions. This causes a final lexical L to remain L in a positively expressed question, while this final L becomes a LH rise in a negatively expressed question: [ùbì] ‘book’ / [ùbí] ‘book? [negative]’. (my boldface and braces)

No example has been found in NEA for these last three types of question marking.

Generally, it seems, the same question marking strategies as in polar questions can also be employed in other question types. However, this has not actually been investigated. König & Siemund (2007: 292), for instance, argue that “alternative questions can be neglected since, at least from our current perspective, they do not seem to show any striking typological variation.” This general negligence of *alternative questions* may be partly due to the fact that in any given language they are known to be much less frequent than polar or content questions (Hoymann 2010: 2728). But Siemund (2001) and König & Siemund (2007) are clearly wrong in their assessment that **alternative questions** do not exhibit any interesting variation to be discovered. On the contrary, they actually exhibit much more variation than polar questions because, in addition to the question marking strategies encountered above, they show interaction with coordination, have two or more possible loci of marking, and display interesting patterns of ellipsis that may affect the question marking.

The simplest marking strategy is a mere juxtaposition of the two alternatives. However, the two alternatives may still be marked with intonation patterns that are not always specified. For instance, in Amis (13) each alternative is marked with “a leveling-rising-falling intonation pattern” (Huang et al. 1999: 650).

- (13) Amis (Nuclear Austronesian, Austronesian)  
*ma-tayal kísu ma-fúti?*  
 AG.FOC-work 2SG.NOM AG.FOC-sleep  
 ‘Are you going to work or sleep?’ (Huang et al. 1999: 651)

So-called A-not-A questions, frequently encountered in MSEA, are perhaps best analyzed as a subtype of this type of alternative question marking with an additional negator.

- (14) Mon (Monic, Austroasiatic)  
*klɜŋ hùʔ klɜŋ?*  
 come NEG come  
 ‘Are (you) coming (or not)?’ (Clark 1985: 60)

In other cases the two alternatives may be conjoined with the help of a disjunction. For example, Saisiyat (15) makes exclusive use of a disjunction, but lacks any further question marking, including intonation.

#### 4 The typology of questions

- (15) Saisiyat (Nuclear Austronesian, Austronesian)

*nifo ?am nyaw a ?am ?ahæ??*

2SG.GEN want cat or want dog

‘Do you want a cat or do you want a dog?’ (Huang et al. 1999: 652)

Some languages such as Finnish have a special interrogative disjunction (*tai*) that is not identical to the standard disjunction (*vai*) (e.g., Haspelmath 2007: 25). In other languages there is no disjunction but a question marker, for example on the first alternative. Consider the following negative alternative question (16), which exhibits the same question marker found in polar questions.

- (16) Guiqiong (Qiangic, Trans-Himalayan)

*zo gutchiŋ dzi lɛ, mɛ-dzi?*

3SG PN COP Q NEG-COP

‘Is (s)he a Guiqiong or not?’ (Jiang Li 2015: 305)

In English (as in the translation of 16 above) the polar question marking strategy on the first alternative (in this case word order change) is combined with a disjunction, which appears to be a common European phenomenon. However, this is combined with a special intonation contour in English, which rises on the first and falls on the second alternative. In other languages, there is a question marker attached to the second alternative. The following example (17) is also a negative alternative question.

- (17) Palula (Dardic, Indo-European)

*tu the phedil-u, ki na?*

2SG.NOM to arrive.PFV-SG.M ?Q NEG

‘Did you receive it or not?’ (Liljegren 2016: 404)

This, again, may be combined with disjunctions. In other languages there are question markers on each alternative, with or without disjunction. Examples for these types can be found below such as in (21). Table 4.4 schematically shows some possible types of interaction of disjunction and question marking. Of course, it is simplified and does not show all possible marking strategies such as the use of intonation, particles, clitics, affixes etc. It merely schematically illustrates juxtaposition, single marking on the first or second alternative and double marking, all of which may combine with disjunctions. It becomes apparent that there are dozens of combinations of these patterns with different marking strategies, which makes it impossible to present them all in this section. Each type, furthermore, can interact with other domains such as negation.

What is more, the plethora of different patterns in Table 4.4 above does not even cover all alternative question marking strategies found in the languages of the world. Khwarshi, for example, in addition to double marking, contains cases in which the disjunction *ya(gi)*, borrowed from Awar, is not employed once but twice (18).

Table 4.4: Schematic interaction of disjunction and question marking.

1st alternative		2nd alternative
X	+/-DISJ	Y
X-Q	+/-DISJ	Y
X	+/-DISJ	Y-Q
X-Q	+/-DISJ	Y-Q

(18) Khwarshi (Tsezic)

*me ya lo-k n-eq-še ya lac'á-k l-i-še?*  
 2SG.ERG or water(IV)-Q IV-bring-PRS or food(IV)-Q IV-do-PRS

‘Will you bring water or make the meal?’ (Khalilova 2009: 458)

The language Edo (Niger-Congo) has a disjunction *rà*, either once between the two alternative, or twice following each alternative (Omoruyi 1988: 23). Additionally, the markers on the different alternatives are not necessarily identical as can be illustrated with data from Tshangla as spoken in Bhutan (19).

(19) Tshangla (Trans-Himalayan)

*ser-ga rengan tang-pe mo, shing-ga rengan tang-pe ya?*  
 gold-LOC ladder bridge-INF Q wood-LOC ladder bridge-INF Q

‘Should I put up a silver ladder or a wooden ladder?’ (Andvik 2010: 193)

In Tshangla, *mo* is also a polar question marker and *ya*, which is optional in alternative questions, is also found in content questions.

In some languages there is a complex expression meaning ‘(and) if not’ (20), which functions more or less like a disjunction but should be kept distinct as it is etymologically transparent.

(20) †Ākhoe Hailom (Khoe)

*uri ra ari-b.a tama-s ga i-o !gû ra ari-b.a?*  
 jump PROG dog-3SG.M NEG-3SG.F POT STAT-if walk PROG dog-3SG.M

‘Does the dog jump or does the dog walk?’ (Hoymann 2010: 2733)

Yet another dimension of variation concerns the number of alternatives. While it is true that the most typical alternative questions exhibit two alternatives, there are also examples with more than two, such as in (21).

(21) Mauwake (Trans-New Guinea)

*no matukar ikiw-i-nan=i dylup=i e sarang?*  
 2SG PN go-N.PST-FUT.2SG=Q PN=Q OR PN

‘Will you go to Matukar, Dylup, or Sarang?’ (Berghäll 2015: 310)

#### 4 The typology of questions

Mauwake usually has an enclitic =*i* at the first alternative and a disjunction before the second. When three alternatives are present, the first two take the enclitic. This example also illustrates that the question markers in the individual alternatives do not have to attach at the same place. When the set of possible answers is expected to be open, the construction differs slightly and the second alternative also takes the question marker (22).

- (22) Mauwake (Trans-New Guinea)  
*matukar ikiw-i-nan=i e dylup ikiw-i-nan=i?*  
 PN go-N.PST-FUT.2SG=Q OR PN go-N.PST-FUT.2SG=Q  
 ‘Will you go to Matukar or Dylup (or perhaps neither)?’ (Berghäll 2015: 311)

Some languages do not allow ellipsis of identical parts (e.g., the Austronesian language Rukai, Zeitoun 2007). All other languages allow some form of deletion. A very useful distinction that was introduced by Huang et al. (1999) for Austronesian languages on Taiwan is that between forward (anapipsis, 23b) and backward deletion (catalipsis, 23c) (see also Haspelmath 2007: 39).

- (23) Mandarin (Trans-Himalayan)
- a. *nǐ qù zhōngguó (háishi) bú qù zhōngguó?*  
 2SG go China or.Q NEG go China
  - b. *nǐ qù zhōngguó (háishi) bú qù — ?*  
 2SG go China or.Q NEG go (China)
  - c. *nǐ qù — (háishi) bú qù zhōngguó?*  
 2SG go (China) or.Q NEG go China
- ‘Are you going to China or are you not going to China?’ (elicited, own knowledge, cf. Hölzl 2015e)

In alternative questions the part that is not focused on may fall victim to ellipsis. In other words, (elliptical) alternative questions are somewhat similar to focus questions. This contrasts with the common assumption of alternative questions being related to polar questions, exclusively (e.g., Siemund 2001).

Content question marking has not been investigated very often. Many languages have morphosyntactically unmarked content questions, but these may exhibit special intonation patterns that often are not clearly specified in the available descriptions. The remaining languages seem to make use of all the most common question marking strategies discussed above for polar questions and will thus be excluded here. Many examples can be found in §5.

The marking of *focus questions* is difficult to investigate because most grammatical descriptions simply do not mention it. Most likely, they can exhibit more or less the same range of marking strategies as polar questions. Given their interaction with the domain of focus, they will be discussed further in §4.2.3 on the interaction of functional domains. Several examples can be found throughout Chapter 5.

Table 4.5: A typology of question tags according to Axelsson (2011: 803)

invariant question tags	- neutral - polarity-biased - polarity-dependent	no restrictions
variant question tags	- lexically-dependent - marginal grammatically- dependent - grammatically- dependent	↑  most restrictions

*Tag questions* have been excluded from the list of central question types in this study. Nevertheless, some information on their formal properties seem to be in order. Perhaps the best typology of tag questions has been given by Axelsson (2011: 803) (Figure 4.5). A main difference is drawn between invariant and variant tags. Invariant tags appear to be more common, both cross-linguistically and in NEA. Each is furthermore divided into three different subtypes.

So-called neutral and polarity-biased question tags are neutral with respect to the polarity of the anchor, although the latter often prefers positive or negative anchors. Polarity-dependent question tags, as the name suggests, are restricted to either positive or negative anchors. Consider the following examples from English (24), where the first is a neutral (non-dependent) and the latter a grammatically-dependent question tag (own knowledge).

(24) English

- a. *You want coffee, **right**?*
- b. *You want coffee, **don't you**?*

Marginal grammatically-dependent question tags, on the other hand, “are cases where the use of a certain question tag is dependent on a certain grammatical feature in the anchor (other than polarity), but where there are no variable grammatical features in the tag itself.” (Axelsson 2011: 805) In lexically-dependent question tags, a lexical element of the anchor is also found in the tag (Axelsson 2011: 805). There are relatively many languages in NEA for which no tag questions are attested. While at least in some cases this may be due to the lack of sufficient information, tag questions most likely are not a universal property of language.

Another useful dimension of question tags that is somewhat less relevant for other question markers concerns its *etymological transparency*. German, for example, has a variety of tags, among which we find a form *ge(lle)* that is completely opaque from a synchronic perspective (25a). German *richtig*, on the other hand, is a common adjective related to English *right* (25b). Both are neutral question tags.

#### 4 The typology of questions

(25) German

- a. *Du magst Kaffee, ge(lle)?*
- b. *Du magst Kaffee, richtig?*

The meaning and word order are identical to the English sentence *You want coffee, right?* above (24a). In fact, most question markers are opaque from a synchronic perspective. Question tags, on the other hand, are frequently transparent. Question markers furthermore tend to be extremely short (see §6.1.1). Question tags certainly can be short as well (e.g., English *eh?*), but generally tend to be longer and more complex than usual question markers (e.g., English *isn't it?*, Mandarin *dui-bu-dui?*). These properties underline their separate status.

Mithun (2012: 2167) roughly differentiates between epistemic (e.g., informational, confirmatory), and affective (e.g., facilitating, attitudinal, peremptory, aggressive) functions of tags. Axelsson (2011) crucially investigated only confirmation seeking (perhaps better called epistemic) question tags, which reduces the problem of their classification considerably. The typology correctly excludes confirmation seeking constructions that are not formally tag questions (Axelsson 2011: 796). Hadiyya (26), for example, has a confirmation seeking suffix *-lla*, which combines with the polar question marker *-nni(yye)*.

(26) Hadiyya (Cushitic, Afroasiatic)

- kaa ii diinate mass-i-t-aa-tto-lla-yyo-nni?*  
2SG.VOC 1SG.GEN money.ACC take-E-2SG-PRS.PFV-2SG-CONF-NEG-Q  
'You have taken my money, haven't you?' (Sulamo 2013: 27)

Given the fact that the question is one single sentence, it is better classified as a special kind of polar question. §4.4 elaborates on the classification of tag questions. Non-epistemic uses are likewise excluded from this study.

#### 4.2.2 The scope of question marking

While different marking strategies for questions are well-known, it is usually not recognized that these differ in their semantic scope over different question types (but among others see Dixon 2012: 389-390 and especially Hölzl 2015e; 2016b). Given the lack of information for NEA, this study will make use of a limited conceptual space shown in Figure 4.1 that only includes the most central question types. As can be seen, polar questions take a central position while other types—especially content questions—have a peripheral position. Solid lines indicate the possibilities that two categories may be marked with the same marker. The semantic scope of a given marker may be shown as a closed line that encloses those categories that may be marked by it (i.e. its semantic scope).

There is one possible implicational universal that needs further testing in other parts of the world but seems reasonably robust for now.

(27) Content questions are only marked in the same way as focus or alternative questions if polar questions are also marked in the same way.

The universal is represented on the conceptual space as the lack of connecting lines between categories (Figure 4.1). These presumably impossible connections are given as dashed lines. Note that this is an example for the so-called *Semantic Map Connectivity Hypothesis*: “any relevant language-specific and/or construction-specific category should map onto a *connected region* in conceptual space” (Croft 2003: 134). A possible counterexample from Tshangla, which allows the use of the content question marker *ya* in alternative questions, can be found in §4.2.1.

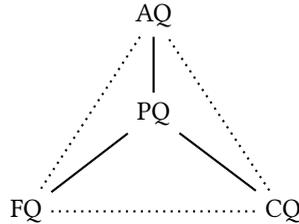


Figure 4.1: Limited conceptual space of question marking

Another possible implicational universal concerns the dashed line between focus and alternative questions (Hölzl 2015e).

- (28) Focus and alternative questions can only be marked in the same way if polar questions are also marked in the same way.

Only one possible exception (the Nyulnyulan language Bardi spoken in Australia, see 3 above) was found within the global 50 language sample investigated by Hölzl (2015e). An obstacle for confirming or disproving the universal is severely hampered by the lack of adequate data for the majority of languages. The dashed lines are also meant to indicate that such connections might be possible after all but clearly are dispreferred.

If the conceptual space is universally applicable, which should be the long-term goal, then it poses several powerful constraints on how markers can expand their scope. An extension of the semantic scope of a given marker, for example, is only possible if there is a connection in conceptual space. Every language shows a distinctive semantic map, but languages may have similar patterns due to universals, tendencies, chance, language contact or common inheritance. Given that question markers are often and freely borrowed from one language to the next, semantic maps easily change their shape.

Content questions, which often remain unmarked morphosyntactically, are a special case. By comparing polar and content questions and further differentiating between morphosyntactically marked versus unmarked content questions, one gets a matrix of four language types (Table 4.6).

Type 4 appears to be the most and Type 3 the least common, cross-linguistically.<sup>1</sup> In sum, there is a deep bifurcation between content questions on the one hand and polar

<sup>1</sup>Given the lack of information Hölzl (2015e) omitted intonation, which should be included in future studies.

#### 4 The typology of questions

Table 4.6: Polar and content question marking strategies among 50 languages, based on Hölzl (2015e); the classification of three languages remained unclear

	CQ marked	CQ unmarked
PQ=CQ	Type 1: 9	Type 3: 1?
PQ≠CQ	Type 2: 9	Type 4: 28

questions on the other. However, as we will see more clearly in Chapter 6, polar questions have closer relations to the other question types.

#### 4.2.3 Interaction of functional domains

The term *functional domain* here covers broad universal categories such as negation, focus, or question marking, which themselves have many subcategories. Hölzl (2016b: 24) distinguished between four different types of interaction between such functional domains shown in (29).

- (29) a. grammaticalization (1)  
 b. combination (2)  
 c. fusion (3)  
 d. interaction (split types) (4)

For practical purposes, the combination of disjunction with question marking was already covered above in §4.2.1.

(1) Grammaticalization in this context is understood as a cover term for the *shift in meaning* of a linguistic element from one functional domain to another. Many details, of course, are language- and construction-specific, but here only a cursory overview similar to the *World Lexicon of Grammaticalization* (Heine & Kuteva 2002) can be given (cf. Hölzl 2015e). Consider the following polar question from a language in Nepal (30).

- (30) Bantawa (Kiranti, Trans-Himalayan)  
*am-k<sup>h</sup>e ham-si ti-k<sup>h</sup>ar-a-ʔo?*  
 2SG.GEN-lice swap-SUP 2AS-GO-PST-Q  
 ‘Did you go to swap lice?’ (i.e. ‘Did you go to have sex?’) (Doornenbal 2009: 205)

The marker *-ʔo* has been glossed as a question marker, but it is really a nominalizer, which is presumably the reason why the example has an additional semantic component ‘is it the fact that’. A similar development has also been described for Tucanoan languages in South America, which

exhibit a historical and semantic relationship between nominalizations and questions. We have also tried to demonstrate that formally the latter originate from

the former through a process of upgrading a nominalized predication to the status of an independent utterance from an inferential or mirative construction. Semantically, the interrogative meaning must have become conventionalized via stages expressing doubt or surprise. (van der Auwera & Idiatov 2008: 46)

Whether exactly the same developmental path was followed in Bantawa or other languages with this phenomenon is not known to me.

Two other well-known examples are the development of disjunctions and negators to polar question markers. However, both of these developments usually start within the context of an elliptic alternative question. In some languages such as Edo (31), the second alternative is fully elliptic and the disjunction can take over the function of a polar question because no second alternative is specified (cf. Dixon 2012: 399-400).

(31) Edo (Niger-Congo)

- a. *Òsàró bọ òwá rà Òsàró rhié òkhùò?*  
 PN build house or PN take woman  
 ‘Did Osaro build a house or did he marry a woman?’
- b. *Òsàró bọ òwá rà?*  
 PN build house Q  
 ‘Did Osaro build a house?’ (Omoruyi 1988: 22, 23)

Similarly, negators can develop into polar question markers in negative alternative questions when the second alternative only consists of the negator. Examples of this sort can be found in Mandarin (§5.9.2.1), for instance. A related development seems to start from negative alternative questions as well, but in this case the first alternative appears to have been deleted. In Kham (Trans-Himalayan), for example, the prefix *ma-* can express both negation and polar questions (Watters 2002: 96-101). Negators such as German *nich(t)* ‘not’ can also develop into question tags.

Yet another frequent development is from interrogatives to polar question markers and question tags. This development is very rare in NEA but many examples can be found in Indo-European languages (§5.5.2, Hackstein 2013: 100). Example (10) from Bengali above, for example, contains the polar question marker *ki*, which is most likely derived from the interrogative *ki* ‘what’ (Thompson 2012: 200-203), see also (17). In the language Palula the interrogative *ga* ‘what’ developed into a question tag (32).

(32) Palula (Dardic, Indo-European)

- so gúum ga?*  
 3SG.NOM go.PFV.SG.M what  
 ‘He left, didn’t he?’ (Liljegren 2016: 404)

This development can also be found in other languages of South Asia. For instance, the Dravidian language Kurux employs the interrogative *ender* ‘what’ as a question marker in sentence-initial position (Kobayashi & Tirkey 2017: 241-242). Another example mentioned above stems from Bardi. §6.1.3 summarizes the most important grammaticalization paths found during this study (see also Bencini 2003).

#### 4 The typology of questions

(2) Question marking is frequently *combined* with interrogatives in content questions and disjunctions in alternative questions. Interrogatives (~ indefinites) are almost universal, but there are many languages without disjunctions, for example in northern NEA. Another special case concerns focus markers that are frequently present in focus and sometimes other question types (Figure 4.2). In English, for example, focus questions are expressed by usual polar question marking and additional intonational focus or a cleft construction (33).

(33) English

- a. Did you go there?
- b. Is it you who went there?

In both cases focus and question marking are merely combined with each other. For practical purposes disjunctions and focus marking will be treated together with question marking in this study, but one should keep in mind that they really belong to different functional domains that merely overlap with each other.

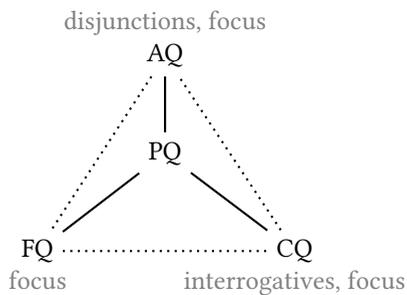


Figure 4.2: Typical interaction of question marking with other functional domains

Previous studies of question marking have presumably focused on polar questions, because these exhibit the least interference with other functional domains.

(3) In instances of *fusion*, on the other hand, a question marker also has additional functions such as focus marking. When a question marker also functions as a focus marker, it usually attaches to the verb in polar questions and to the focal element in focus questions. Such an example can be found in the South American language Quechua as spoken in Cusco (34).

(34) Cusco Quechua (Quechuan)

- a. *wasi-y-mañ hamu-ηki=chu?*  
 house-1-DAT come-2=Q.FOC  
 ‘Do you come to my house?’

- b. *wasi-y-maŋ=chu hamu-ŋki?*  
 house-1-DAT=Q.FOC come-2  
 ‘Do you come to *my* house?’ (Ebina 2011: 29)

See §6.1.3 for a list of examples from NEA.

(4) The most complex question marking systems are *split systems*. In such languages the choice between different question markers depends on other domains such as person, number, tense, aspect, mood, evidentiality, clause type etc. §6.1.3 lists all instances found in NEA. A relatively simple example can be found in the language Qiang (35), which has a split based on person.

(35) Qiang (Qiangic, Trans-Himalayan)

- a. *?ũ zme ŋua-n-a?*  
 2SG PN COP-2SG-Q  
 ‘Are you a Qiang?’
- b. *the: zme ŋua-Ø-ŋua?*  
 3SG PN COP-(3SG)-Q  
 ‘Is (s)he a Qiang?’ (LaPolla & Huang Chenglong 2003: 180)

Only second person singular forms take the marker *-a* instead of *-ŋua*. Many examples of split types exhibit instances of fusion, but this is not necessarily so, as this example illustrates. An example for a split in combination with fusion stems from the Amazonian language Kulina (36), which combines question marking with gender.

(36) Kulina (Arawan)

- a. *osonaa=ko?*  
 PN=Q.M  
 ‘Is he a Kashinawa?’
- b. *osonaa=ki?*  
 PN=Q.F  
 ‘Is she a Kashinawa?’ (Dienst 2014: 193)

The markers appear in both polar and focus questions. Omotic languages (Afroasiatic) exhibit some of the most complex split systems (see Amha 2007; 2012; Hellenthal 2010: 401ff.; Köhler 2013; 2016; Treis 2014; Hölzl 2016b: 26 and references therein). Again, see §6.1.3 for those split types encountered in NEA.

#### 4.2.4 The number of markers

A dimension not mentioned in Hölzl (2016b) is the sheer amount of question markers present in a given language. Arguably, this is yet another dimension of the complexity of the grammar of questions. There is a certain connection with both the scope of question marking and the interaction with other functional domains. A smaller scope of question

#### 4 The typology of questions

markers is usually, but not necessarily, correlated with a higher number of markers. The question marker =*Ku* in the Tungusic language Evenki, for example, has a broad scope that covers polar, focus, and alternative questions, and, indeed, Evenki has only a rather small amount of other question markers that also depend on the dialect, however. If question marking interacts with certain other domains such as person marking, there tends to be a higher number of markers. The average number and possible variation among the languages of the world is not entirely certain but presumably most languages have at least one or a few question markers. It is, furthermore, not self-explanatory how question markers should be counted at all. For instance, the Tungusic language Manchu has a question marker =*ni* that fuses with certain words such as the negative existential *akū* to yield a complex form *akūn*. Should =*ni* and *-n* be counted as one or two markers? Despite such problems, it is usually unproblematic to establish whether a certain language exhibits a larger or smaller amount of markers relative to other languages. The Nicobarese (Austroasiatic) language Muöt, for example, according to Rajasingh (2014: 114), only has one question marker, namely final rising intonation. The Cushitic (Afro-Asiatic) language Hadiyya, to give a slightly more complex example, has three main question markers, rising polar question intonation, the polar and alternative question marker *-nni(yye)*, and the confirmation seeking suffix *-lla* that is usually combined with *-nni(yye)* (Sulamo 2013). The majority of languages in NEA and worldwide seem to cluster somewhere around this relatively small amount of question markers, but there are some languages with extremely complex question marking systems (e.g., §5.4.2 on Yupik, §5.9.2.1 on Sinitic, and §5.14.2 on Yukaghiric). Perhaps the upper end is formed by Omotic (Afroasiatic) languages, which sometimes exhibit a plethora of several dozen different forms organized in many different paradigms (e.g., Amha 2007; 2012; Hellenthal 2010: 401ff.; Köhler 2013; 2016; Treis 2014; Hölzl 2016b: 26, and references therein).

### 4.3 Interrogatives

What will simply be called *interrogatives* here has variously been termed *wh-words*, *interrogative pronouns*, *interrogative words*, *question words* etc. But these terms are problematic from several perspectives. First, *wh* refers to an English language writing convention exclusively (variously pronounced /h/ ~ /w/), even fails to capture English forms such as *how*, and has no validity whatsoever from a typological perspective. Interrogatives are, furthermore, not necessarily pronominal. Instead, they represent what has been called “a meta word-class, spanning a number of major classes” (Dixon 2002: 80) or “a pan-basic-word-classes word class” (Dixon 2012: 409). As we will see during this section, there are interrogative nouns, verbs, adjectives, adverbs etc. The terms *question word* or *interrogative word* are, therefore, more adequate than the other terms but still problematic. While *pronoun* suggests a connection with grammar, *word* clearly indicates a lexical category. It has been shown by Diessel (2003), however, that interrogatives (and perhaps demonstratives), do not clearly belong to either of these categories. Diessel (2003: 636) is certainly right in his view (also accepted by Cysouw & Hackstein 2011) that

while grammatical markers are commonly derived from lexical expressions, demonstratives and interrogatives cannot be traced back to lexical items. While both are often reinforced by other lexemes, there is no evidence from any language that a new demonstrative or interrogative developed from a lexical source (unless the lexical source first functioned to reinforce a genuine demonstrative or interrogative). All this suggests that demonstratives and interrogatives have a special status in language and should be kept separate from genuine grammatical markers.

Like lexical items, both demonstratives and interrogatives are often the source for several grammatical items. In a brief discussion on Funknet, Heine & Kuteva (p.c. 2018) made me aware of the fact that there are indeed several examples of demonstratives with lexical origins. Nevertheless, interrogatives and perhaps demonstratives might still form a class by themselves that is neither, strictly speaking, lexical, nor grammatical. “Grammatical markers organize the information flow in the ongoing discourse, whereas basic demonstratives and interrogatives are immediately concerned with the speaker-hearer interaction.” (Diessel 2003: 635) Interestingly, the two often share paradigmatic similarities (see below) and the only known language without interrogatives, the Chapacuran language Wari’, uses demonstratives instead (Everett & Kern 2007).

There are several imaginable typologies for interrogatives, but many of them do not make too much sense from a cross-linguistic perspective. For example, one might count the *number of forms* that may be encountered in one language. The number of interrogatives among languages is highly variable. There are none in Wari’ (Everett & Kern 2007) but up to about 30 in German according to my count, including derived forms. However, apart from the practical problem that almost no grammatical description mentions more than a handful of forms, it is by no means clear how such forms should actually be counted. Mackenzie (2009: 1133), for instance, counts “only the simple forms as true interrogative forms”. Similarly, Hengeveld et al. (2012: 46) only include “basic question words”. The necessary condition for these claims is a clear-cut boundary between forms that can be analyzed and those that cannot. However, the existence of such a boundary is far from clear because *analyzability* is clearly “a matter of degree” (Langacker 2008: 352). Let me illustrate this with the help of interrogatives in the Tungusic language Manchu (§5.10.3). There certainly are some non-analyzable “basic” interrogatives such as *we* ‘who’ that even historically are not transparent. Then there are forms such as *atanggi* ‘when’, which is not analyzable synchronically but shares a resonance (a submorpheme) *a~* with several other forms. In all likelihood it is ultimately based on the interrogative *ai* ‘what’, but the derivation remains unclear, since a word meaning perhaps ‘time’ with this form is not attested. Mackenzie (2009) and Hengeveld et al. (2012) would perhaps include both of these forms into the category of “basic question words”, but this is an arbitrary decision. Manchu furthermore has a form *aiseme* ‘why’ that clearly is a combination of *ai* ‘what’ and the quotative *seme*, which in turn may be analyzed as *se-me* ‘say-CVB.IPFV’. Despite its formal analyzability, the semantic side is not fully compositional. Further problems for an analysis are so-called cranberry morphs such as the second element in Manchu *ai-bi-de* ‘where’, which stands opposed to the fully-analyzable form *ai-ba-de* ‘what-place-LOC’. Manchu simply has no independent form *bi* that would explain the sec-

#### 4 The typology of questions

ond element in *ai-bi-de*. It is not the first person singular nominative *bi*, nor the copula *bi* which cannot take any case markers. The most likely scenario is an idiosyncratic development from *ba* ‘place’. In any case, the point is that there are partly analyzable forms that constitute a scale between non-analyzable and fully-analyzable forms (see also Cysouw 2005). This background also means that reconstructions of clear interrogative “stems” for any given proto-language in most instances must in principle be considered problematic. While the analyzability of interrogatives tends to decrease over the course of time if no new forms are built, there may also be a development in the opposite direction, as witnessed in the reanalysis of German *wor.um* ‘around where, about what’ as *wo.rum*, which allows a reconnection to the word *wo* ‘where’ that historically lost the final *-r* (PIE *\*k<sup>w</sup>ór*) and the creation of a new form *rum*.

There are several more possible dimensions for a typology of interrogatives. Some investigations (Heine et al. 1991; Peyraube & Wu 2005; Mackenzie 2009; Hengeveld et al. 2012) have combined several of these dimensions (e.g., analyzability, polysemy, length) into one typology. However, the results that take the form of a hierarchy are simply not valid from a cross-linguistic perspective (Hölzl 2015c). A study mostly neglected in later typologies (but see Peyraube & Wu 2005) has been conducted by Heine et al. (1991), who investigated what they called “metaphorical relations” and how they related to interrogatives in 14 different languages. Their result is a hierarchy that has the following form (37, slightly adapted):

(37) PERSON < THING < ACTIVITY < PLACE < TIME < MANNER < PURPOSE/CAUSE

According to their study, the first four categories on the hierarchy showed minimal phonological and morphological complexity and were often monosyllabic. TIME and QUALITY were slightly more complex. PURPOSE and CAUSE were found to be much more complex and often had the form “what-case”. Furthermore, in the languages investigated, THING and ACTIVITY were claimed not to be differentiated (e.g., English *what*, (*to do what*)). They had several interesting conclusions such as the following:

While it remains unclear what the exact correlations between the linguistic and the cognitive structure of pronouns are, a few assumptions may be tentatively formulated. First, the relative degree of morphological complexity that a pronoun exhibits is likely to correlate to some extent with the relative degree of its **cognitive complexity**. [...] Second, formal similarity between different pronominal categories may be indicative of some kind of conceptual relation between these categories. (Heine et al. 1991: 59, my boldface)

Let us now address an interesting typology by Mackenzie (2009), who, strangely, did not mention the study by Heine et al. (1991). He investigated interrogatives in a sample of 50 languages. More specifically, he concentrated on so-called “cognitive complexity”, which may be accessed through an investigation of system complexity (extent of polysemy), item complexity (extent of analyzability), and signal complexity (number of phonemes, length), all of which were also included by Heine et al. (1991). The result of

his study also takes the form of a hierarchy which has the following form (38, put into a comparable format):

(38) PERSON/THING < PLACE < TIME < MANNER < QUANTITY < CAUSE

The major difference is that Heine et al. (1991) included ACTIVITY instead of QUANTITY. In fact, Mackenzie (2009: 1150) himself noted that “none of the central hypotheses has been fully vindicated”. In my eyes, the main problem is the combination of different typological dimensions that are not directly connected (such as polysemy and length) and thus simply lead to inconclusive results. Mackenzie (2009) furthermore made some minor but unimportant mistakes such as counting letters instead of phonemes for Mandarin and including expressions about the time of day into the category of TIME.

A follow-up study of Mackenzie (2009) was conducted by Hengeveld et al. (2012), who proposed a hierarchy based on “basic question words” (i.e., non-analyzable interrogatives).

(39) PERSON/THING < PLACE < MANNER < QUANTITY/TIME/REASON

Mackenzie (2009), who also investigated this problematic category, found the following slightly deviating hierarchy:

(40) PERSON/THING < PLACE < QUANTITY < MANNER < TIME < REASON

However, the idea of a cross-linguistically valid hierarchy of “basic question words” has to be refuted, too (Hölzl 2015c). For example, Tungusic data result in the hierarchy shown in (41) (see §5.10.3):

(41) PERSON/MANNER/QUANTITY < TIME < THING/REASON < PLACE

As can be seen, there are severe problems such as the completely different location of PLACE on the hierarchy. In other words, such a hierarchy simply does not make sense from a cross-linguistic perspective. There is no reason to assume that a one-dimensional construct is capable of capturing the much more complex phenomenon of interrogatives. There might be some exceptions such as the frequency of certain interrogatives across languages that could converge to a certain degree, but this has not been investigated and turned out to be impossible to investigate for NEA due to lack of sufficient data for almost all languages. It is also possible to investigate the mere length of interrogatives (e.g., German *wer* ‘who’ is shorter than *warum* ‘why’), but there does not appear to be a universal hierarchy either (Hölzl 2015c). At least there may be a tendency for some categories (e.g., ‘who’, ‘what’) to have shorter forms than others (Mackenzie 2009: 1139), but this is not exclusively connected with the overall frequency in texts. For instance, the shortest interrogative in the Tungusic language Nanai is *ui* ‘who’, which is much less frequent than *xooni* ‘how’ (Kazama 2007: 320). Furthermore, there may be some convergence in the order in which interrogatives are learned by children during language acquisition. Previous research indicates a hierarchy of the following sort (Tomasello 2003: 159 and references therein, 42, slightly adapted).

#### 4 The typology of questions

(42) THING/PLACE < PERSON < MANNER/REASON < TIME

However, the hierarchy is based on only a handful of languages and there is insufficient data for most languages in NEA.

The following will address the (1) semantic scope (§4.3.1), (2) word class membership (§4.3.2), (3) diachrony (§4.3.3), (4) inflectional properties (§4.3.4), and (5) the connection of interrogatives to demonstratives (§4.3.5), which, for the purposes of this study, seem to be the most important dimensions for a typology.

##### 4.3.1 Semantic scope of interrogatives

For the illustration of differences in the *semantic scope* of interrogatives consider example 43 from the language Kusunda, a language without clear affiliation spoken in Nepal, and their English translations.

(43) Kusunda

a. *nəti na?*

INT this.AN

‘Who is this?’

b. *nəti ta?*

INT this.INAN

‘What is this?’ (Watters 2006: 48)

The two categories of PERSON and THING are expressed with two different interrogatives (*who* and *what*) in English but with one (*nəti*) in Kusunda. Thus, there is a difference in semantic scope of the interrogatives over different semantic categories. Usually, a narrow semantic scope goes along with a larger number of interrogatives and *vice versa*. In these examples, animateness in Kusunda is expressed by the demonstratives instead. As Cysouw (2005; 2007) has shown, this particular polysemy (PERSON=THING) is rare worldwide but relatively common in South America. In Eurasia it can also be found in Baltic and Tocharian B.

The determination of the semantic scope of a given interrogative presupposes a fixed set of *semantic categories*. However, there is a certain dispute as to how many different categories should be postulated. The comparison in Table 4.7 is not exhaustive, but sufficient for our purposes (see also Mushin 1995 etc.).

There is no agreement in terminology or number of different categories. This study follows Cysouw’s (2005) approach but adds additional categories. Strangely, only Heine et al. (1991) include the categories of ACTIVITY and PURPOSE, of which at least the first is rather crucial from a cross-linguistic perspective, and only Diessel (2003) mentions spatial interrogatives with an allative or ablative meaning. There are, furthermore, many more categories that are not included in the list, but the most prominent ones are certainly represented. One category that should perhaps be added is KIND, which might

Table 4.7: A selection of different categorizations of interrogatives (Hölzl 2015c)

Muysken & Smith 1990	Heine et al. 1991	Diessel 2003	Cysouw 2005	Mackenzie 2009	Dixon 2012
who	PERSON	person	PERSON	?individual [+/-hum]	who
what	OBJECT	thing	THING		what
-	ACTIVITY	-	-	-	-
which	-	-	SELECTION	-	which
-	-	amount	QUANTITY	quantity	how many/ how much
why	CAUSE	-	REASON	reason	why
-	PURPOSE	-	-	-	-
how	QUALITY	manner	MANNER	manner	how
where	SPACE	place	PLACE	location	where
-	-	direction:to	-	-	-
-	-	direction:from	-	-	-
when	TIME	time	TIME	?time	when

have been overlooked because English *what kind of* and similar forms in other European languages is fully analyzable and thus appears non-basic. Nevertheless, this category has to be distinguished from the category of SELECTION, e.g. English *which (one)*, which does not classify but individualizes a given referent. Thus, this study tentatively distinguishes the categories of PERSON, THING, SELECTION, ACTIVITY, CAUSE, MANNER, QUANTITY, PLACE, TIME, and KIND. Some of these have secondary subcategories such as COUNT (*how many*) or MASS (*how much*) in QUANTITY and LOCATION (*where*), DIRECTION (*whither*), and SOURCE (*whence*) in PLACE. The category PURPOSE will not be distinguished from CAUSE as it does not appear to play a crucial role for languages in NEA. The same is true for the difference between MANNER and QUALITY. There are some additional categories, but including them is not absolutely necessary because only a handful of forms is attested for most languages in NEA. There are a number of subcategories that will not be addressed any further. Pite Saami (Uralic), for example, apart from the selective interrogative *mikir-* ‘which’ has a special interrogative *gãb-* ‘which one (out of two) (SG), which two (PL)’ (Wilbur 2014: 123).

The question ‘What is your name?’ (see Idiatov 2007) often allows the use of two different interrogatives, ‘who’ and ‘what’. In some languages (e.g., 44) both interrogatives may be used.

(44) Abui (Timor-Alor-Pantar)

- a. *a-ne*                    *nala?*  
 2SG.INAL-name what  
 ‘What is your name?’

#### 4 The typology of questions

- b. *a-ne*                      *maa?*  
 2SG.INAL-name who  
 ‘What is your name?’ (Kratochvíl 2007: 129)

Thus, there is no absolutely clear-cut or at least a language specific boundary between the categories of PERSON and THING. Similar problems exist for other categories such as MASS versus COUNT.

Interrogatives have what can be called *schematic* (e.g., Langacker 2008) meaning and they express basic semantic categories (e.g., Schulze 2007). Direct evidence for the basic meaning of interrogatives can be found in many languages that have *transparent* interrogatives (Muysken & Smith 1990) such as English *what kind of* or *what for*. A list of frequent elements that are combined with interrogatives can be found in Table 4.8. For example, the Trans-Himalayan language Anong has a rather general interrogative  $k^h a^{55}$  ~  $k^h a^{31}$  that, if combined with a personal classifier, forms the interrogative  $k^h a^{31}-io^{55}$  ‘who’ (Sun Hongkai et al. 2009: 73-74). In Sheko (Omotic, Afroasiatic) the interrogative *yírà* ‘what’ can take a “motive” marker; the resulting form *yír-ěfñtà* has acquired the meaning ‘why’ (Hellenthal 2010: 411-412). Useful but much less common alternatives for the designation of interrogatives are *epistememes* (Mushin 1995) or *ignoratives* (Miyaoka 2012: 443-461), which both emphasize their relation to knowledge.

Table 4.8: Examples for semantic connections between interrogatives and basic nouns etc.; see Chapter 5 for many examples

Category	English	Basic Elements
PERSON	who	man, person, one, DEM, CLF
THING	what	thing
SELECTION	which (one)	one, CLF
KIND	what kind of	kind, sort, class
ACTIVITY	to do what	to do, to make
CAUSE	why, what for	cause, reason, DAT, CVB, PURP
MANNER	how	way, fashion, manner
QUANTITY, MASS	how much	much, few, ?amount
QUANTITY, COUNT	how many	many, ?number
PLACE, LOCATION	where	place, side, LOC
PLACE, DIRECTION	whither, where to	direction, ALL
PLACE, SOURCE	whence, where from	?source, ABL
TIME	when	time (+ LOC)

Figure 4.3 is a slightly revised version of Cysouw’s (2005) illustration of major pathways of the derivation of interrogatives, and may also be understood as a conceptual space for interrogatives (Hölzl 2015c). Similar to the conceptual space for question marking in §4.2.2, this conceptual space of interrogatives allows a comparison of the semantic scope of individual interrogatives within one or across several languages. Connections

between categories indicate the possibility that they can be covered by the same interrogative. Arrows furthermore show common paths of developments, either merely semantic or by means of derivation and inflection.

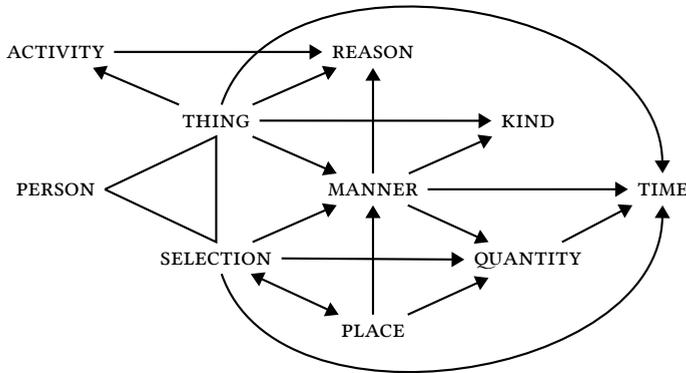


Figure 4.3: A conceptual space of interrogatives

Cross-linguistic data suggest that interrogatives meaning ‘what’ or ‘which’ are the unmarked and most basic members of the interrogative system and often serve as the basis for the derivation of other interrogatives. The grammaticalization of interrogatives to question markers and the use of interrogatives in open alternative questions offer additional evidence for this hypothesis; in both cases it is typically an unmarked interrogative with the meaning ‘what’ that is employed. The category PERSON occupies a special position as it appears to be less prone to changes and more stable diachronically.

The conceptual space was in need of several slight revisions. For NEA the category of ACTIVITY had to be added; it is integrated into the map with the following connections: THING→ACTIVITY→REASON (Hölzl 2015c). An example is Manchu, which has an interrogative *ai* ‘what’. This interrogative may take a verbalizer *-na-* to yield *ai-na-* ‘to do what’, which, in turn, may take the imperfective converb marker *-me*, resulting in the complex interrogative *ai-na-me* ‘why’ (literally ‘doing what’ or ‘in order to do what’). The category of KIND has also been tentatively added. For example, English *what kind of* suggests a connection THING→KIND (see also Idiatov 2007: 51ff.) and Mandarin *zěnme yàng de* MANNER→KIND (*zěnme* ‘how’, *yàng* ‘kind, type’, *de* ‘ATTR’). It may be necessary to update further aspects of the conceptual space in future studies such as a possible connection SELECTION→KIND, but for NEA the most important aspects are present.

Apart from these categories, the space also lacks the categories of DIRECTION and SOURCE, that are clearly related to the category of PLACE. Further categories such as translatives or prolatives will be ignored due to a lack of data for most languages in NEA. Figure 4.4 shows these three categories on a small conceptual space (Hölzl 2015c) that is already known from studies in case marking (e.g., Creissels 2006). Within Cysouw’s conceptual space, the category of PLACE appears to cover not only LOCATION but the two categories of DIRECTION and SOURCE as well. For instance, Manchu *absi* ‘how’ derives

#### 4 The typology of questions

from a form meaning ‘whither’ and German *woher* ‘whence’ may also mean ‘how, why’ in certain contexts (e.g., *woher denn?* ‘why then?’). The conceptual space for locative interrogatives may be conceptualized as the result of zooming in on the category of PLACE. A close-up examination of QUANTITY reveals the limited conceptual space MASS–COUNT.

Within the conceptual space for locative interrogatives, languages differ with respect to scope, markedness, whether they have case marking or special forms, and whether case markers are also found on nouns or not. English used to have the special forms *whence* and *whither*, but they have been replaced with the case marked forms *where to* and *where from*. Within the new system, *where* is unmarked for case. While in English *to* and *from* are usual case markers (or prepositions), German *wo-hin* and *wo-her* (derived from *wo* ‘where’) have special suffixes that may otherwise only be found in the demonstratives (and as verboids, see §5.5.3.2). English and German have three different forms, but Italian *dove* has scope over both LOCATION (*Dove sei?* ‘Where are you?’) and DIRECTION (*Dove vai?* ‘Where are you going?’), while SOURCE is expressed with *di/da dove* (*Di dove sei?* ‘Where are you from?’, *Da dove vieni?* ‘Where are you coming from?’). A recent book on spatial interrogatives that appeared after finishing this book could unfortunately not be taken into account (Stolz et al. 2017).

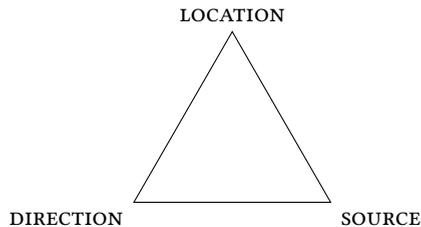


Figure 4.4: A simplified conceptual space for subcategories of PLACE

#### 4.3.2 Word class membership of interrogatives

Typical word class membership of interrogatives is relatively straightforward (Table 4.9), although there is some cross-linguistic variation. As mentioned before, interrogatives belong to a lot of different word classes. There are several clues for determining the word class of a certain interrogative such as inflectional properties or open derivations. For instance, interrogative verbs in many languages are either combinations of the interrogative ‘what’ with a plain verb such as ‘to do’ (English *to do what*) or contain a verbalizing element (Manchu *ai-na-* ‘what-v-’). To take another example, causal interrogatives are often verbs with a converb marker (Even *ja-mi* ‘why’) or nouns with a case marker such as the dative (Buryat *yüün-de* ‘why’). Nevertheless, converb and case markers are often related with each other diachronically and fulfill similar adverbial functions. See Chapter 5 for many more examples.

Paradigms in the Australian language Djabugay (Pama-Nyungan), to give but one additional example, show an interesting split between a pronominal accusative pattern on the one hand (PERSON) and a nominal ergative marking on the other (THING) (Table 4.10).

Table 4.9: Typical word class membership of different interrogatives

Category	English	Typical word class
PERSON	who	pronoun, noun
THING	what	noun, pronoun
SELECTION	which (one)	adjective, (pro)noun
KIND	what kind of	adjective
ACTIVITY	to do what	verb
CAUSE	why, what for	adverb
MANNER	how	adverb
QUANTITY	how many/much	adjective, numeral
PLACE	where, whither, whence	adverb
TIME	whither, where to	adverb
DIRECTION	whence, where from	adverb
SOURCE	when	adverb

Table 4.10: Inflection of Djabugay (Pama-Nyungan) interrogatives (Nau 1999: 135)

	PERSON	THING
S	<b>dju:</b>	nyirrangu
A	<b>dju:</b>	<b>nyi:</b>
O	dju:ny	<b>nyi:</b>

### 4.3.3 The diachrony of interrogatives

The diachrony of interrogatives can be described with a limited set of developmental paths summarized in Table 4.11. (A) Interrogatives may simply be too old to be analyzable at all. To repeat the example from Chapter 1, English *where* or German *wo(r-)* go back directly to Proto-Indo-European *\*k<sup>w</sup>ór*. Apart from phonological changes, the form has been preserved over the course of several millennia. A special subtype of this is the loss of the resonance, i.e. the existence of the same initial sounds in several interrogatives (Bickel & Nichols 2007; Mackenzie 2009, Chapter 1). Such a resonance is usually the sign of an old etymological connection between the participating interrogatives. Given the predominance of suffixes over prefixes (e.g., Manchu *ai-de* ‘what-DAT’) and the dominant

#### 4 The typology of questions

word order IntN (e.g., Manchu *ai-ba-* ‘what-place-’) etc., this feature might be especially pronounced in NEA. Phonological changes, such as the bonding and fusion of such analyzable forms, lead to the emergence of resonances. In most Tungusic languages, an original resonance that is preserved in some languages such as Nanai *x-*, was lost completely (e.g., Nanai *xai* vs. Manchu *ai* ‘what’). Such a development is unique as it affects all interrogatives that share the changing phonological feature. All other changes affect only one or two interrogatives at once.

Table 4.11: The diachrony of interrogatives excluding developments from interrogatives to other domains (Hölzl 2015c); PT = Proto-Tungusic

	Schematic	Details	Example
A	INT <sub>1</sub> > INT <sub>1</sub>	phonological changes	PIE *k <sup>w</sup> ód > OE <i>hwæt</i> > NE <i>what</i>
B	INT <sub>1</sub> > INT <sub>2</sub>	semantic changes	Wutun <i>age</i> ‘which (one) > who’
C	INT <sub>1</sub> -X <sub>GRAM</sub> > INT <sub>2</sub>	inflection (> fusion)	English <i>where to</i>
D	INT <sub>1</sub> -X <sub>LEX</sub> > INT <sub>2</sub>	derivation, reinforcement (> fusion)	English <i>how much</i>
E	(INT <sub>1</sub> )-X <sub>LEX</sub> > INT <sub>2</sub>	replacement	Italian <i>che</i> > <i>che cosa</i> > <i>cosa</i> ‘what’
F	INT <sub>1</sub> , INT <sub>2</sub> > INT <sub>3</sub>	convergence	PT * <i>ja</i> , * <i>Kai</i> > Kh. Evenki <i>i(i)-</i>
G	?X <sub>LEX</sub> > INT	?grammaticalization	?Evenki <i>anjii</i> ‘thing > INT’

(B) There may be semantic changes that leave the formal side more or less intact or are at least not directly connected with it. One such change is the development from the meaning ‘which one’ to ‘who’ as it can be found in several languages in NEA such as the Sinitic language Wutun (see also Idiatov 2007). Both demonstratives and interrogatives are frequently reinforced with the help of other elements, (C) grammatical (e.g., Manchu *ai-de* ‘what-LOC > where, why, how’) or (D) lexical (Manchu *ai-ba-(de)* ‘what-place-(LOC) > where’). Over the course of time these two elements normally fuse into one form. Possible developments of these last three types can also be found on Cysouw’s (2005) conceptual space (Figure 4.4). (E) In some instances, however, the original interrogative may be dropped such as in Italian (*che*) *cosa* ‘thing > what’. This is somewhat reminiscent of one of the well-known Jespersen cycles for negation such as the gradual replacement of *ne* by *pas* in French. (F) Convergence is very rare and within NEA seems to be restricted to Tungusic languages. In some languages such as Khamnigan Evenki, perhaps due to phonological changes, two different interrogative stems merged into one form. This might be treated as a subtype of change (A) but has an impact on both the form and function of several interrogatives. (G) Whether lexical items can directly develop into interrogatives as argued by Schulze (2007), for instance, is highly disputed. Most scholars deny this possibility altogether (e.g., Diessel 2003; Cysouw & Hackstein 2011) and I

tend to agree. There may be some valid examples, such as in Evenki (see §5.10.3), but this certainly is much less common than developments (C), (D), and even (E).

Most of these changes have been taken into account by Muysken & Smith (1990), who developed one of the best typologies of interrogative systems (Table 4.12).

Table 4.12: The typology of interrogatives according to Muysken & Smith (1990)

	Chinese Pidgin English	Sranan	Jamaican	Latin	KiNubi
who	who(-man)	(o)s(u)ma	huu-dat	quis	munú
what	wat ting	(o)san	wa(t)/we/wara	quid	s(h)unú
when	wat-time	o-ten	wen(-taym)	quando	mitéén
where	wat-side	(o)pe	we(-paat)	cur	wén
type	transparent	atrophied	mixed transparent	fused	opaque
	simple	→	→	→	complex

Muysken & Smith (1990) differentiated five different types of interrogative systems. Analyzable combinations of interrogatives with other elements are called *transparent*. The fusion of such analyzable forms leads to fused systems such as in Latin, which in most forms are still related but synchronically not analyzable. The system in KiNubi does not even exhibit such a relic and can be called *opaque*, as the interrogatives are synchronically non-analyzable. Jamaican Creole has both analyzable forms such as *huu-dat* (< English *who-that*) or *we(-paat)* (< English *where-part*), and non-analyzable forms such as *wa(t)* (< English *what*) and therefore can be called *mixed-transparent*. Quite rare are atrophied interrogative systems that used to be transparent but subsequently lost the actual interrogative marker, as in Italian (*che*) *cosa*. The analyzability of forms, of course, does tend to decrease over the course of time, unless new forms are built. But there may also be a development in the opposite direction, as witnessed in the reanalysis of *wor.um* ‘around where, about what’ as *wo.rum* in German which allows a connection to the word *wo* ‘where’ that historically lost the final *-r* (PIE \**k<sup>w</sup>ór*).

Under extreme contact situations an interrogative system may be disturbed or innovated. Bickerton (2016 [1981]: 65-66) and Muysken & Smith (1990) claim that creole and pidgin languages tend to have transparent interrogative systems. Chapter 1 has argued that this phenomenon might not be restricted to creoles, but could be a more general tendency of simplification due to non-native L2 acquisition of a given language (McWhorter 2007; Trudgill 2011; Operstein 2015). Simplification in this case means the reduction in the number of actual interrogatives, the “regularization of irregularities”, and the “increase in morphological transparency” (Trudgill 2011: 62). For this reason Table 4.12 contains a rough scale of complexity. In most cases, innovative interrogative systems are based on an interrogative meaning ‘what’ or ‘which’. An exception to this rule is the language Pichis Ashéninca as described by Cysouw (2007), in which this function is fulfilled by an interrogative meaning ‘where’ (see also §5.5.3.2 on German).

### 4.3.4 Inflectional properties of interrogatives

The inflectional properties of interrogatives are often quite complex and can only be briefly sketched here (see Mushin 1995; Nau 1999; Siemund 2001: 1020–1023 among others). Chapter 5 gives a great many examples of inflected interrogatives.

For the inflection of interrogatives all kinds of morphological types and means are attested cross-linguistically. In Anong (Trans-Himalayan), for instance, the plural of  $k^h a^{31} i o^{55}$  ‘who (SG)’ is formed by reduplication:  $k^h a^{31} i o^{55} k^h a^{31} i o^{55}$  ‘who (PL)’ (Sun Hongkai et al. 2009: 74). As seen in §4.3.3, inflected interrogatives often grammaticalize into interrogatives with a different meaning. Locative interrogatives in Anong, for example, exhibit a locative marker that is analyzed as suffix here,  $k^h a^{31} -a^{55}$  ‘which-LOC’ (Sun Hongkai et al. 2009: 73ff.).

Table 4.13: The inflection of interrogatives in Pite Saami (Uralic; Wilbur 2014: 120–121)

	who		what	
	SG	PL	SG	PL
NOM	ge	ge	mij	ma(-h)
GEN	ge-n	ge-j	ma-n	me-j
ACC	ge-v	ge-jd	ma-v	me-jd ~ ma-jd
ILL	ge-sa	ge-jda	ma-sa	me-jda
INESS	ge-nne	ge-jdne	ma-nne	ma-jdne
ELAT	ge-sste	ge-jsste	ma-sste	ma-jsste
COM	ge-jna	ge-j	ma-jna	me-j

Inflection encompasses verbal (e.g., tense, aspect), nominal (e.g., person, number, gender), and other categories. The inflection of individual interrogatives usually depends on the word class (§4.3.2) and often only a subset of the interrogatives takes inflection. In German, for example, *wer* ‘who’, but not *was* ‘what’, can take morphological case marking. Only the interrogative *wie viel-* ‘how many’ can take the ordinal suffix *-te* that is specific to numerals, e.g. *der wie-viel-te* ‘the how manyth’. The most important inflectional categories for NEA are perhaps number and case that are often organized into paradigms as in Table 4.13.

Interrogatives may express additional nominal categories such as gender (e.g., Icelandic *hver* ‘who.SG.M, who.SG.F’ or *hvað* ‘who.SG.N’, Siemund 2001: 1021), but this plays no important role for most of NEA.

Inflectional properties of interrogatives can often be related to (pro)nouns or verbs, but not necessarily so. Often there is an overlap with the inflection of demonstratives. Consider the paradigms of nouns, demonstratives, and interrogatives of in Pite Saami (Uralic) given in Table 4.14.

In this language there is a strong overlap of the three different paradigms, which nevertheless all have their special properties. Overall the paradigms of the demonstratives

Table 4.14: The inflection of nouns, demonstratives, and interrogatives (PERSON, THING) in Pite Saami, excluding abessive and essive markers for nouns (Wilbur 2014: 93, 116, 120-121)

	N		DEM		INT	
	SG	PL	SG	PL	SG	PL
NOM	-	(-h)	-t	(-h)	-(mij)	(-h)
GEN	(-h)	-j	-n	-j	-n	-j
ACC	-v	-jt	-v	-jt	-v	-jd
ILL	-j	-jda	-sa	-jda	-sa	-jda
INESS	-n	-jn	-n	-jtne	-nne	-jdne
ELAT	-st	-jst	-sste	-jste	-sste	-jsste
COM	-jn(a)	-jn	-jna	-j	-jna	-j

and interrogatives are particularly similar to each other (e.g., GEN.SG *-n* instead of *-h*).

### 4.3.5 Interrogatives and demonstratives

Of the connections to other categories, it is especially demonstratives that will play an important role within this study (§4.3.4, Chapter 5). In fact, many of the typological dimensions mentioned above, such as the diachronic developments, seem to hold for both categories. A connection between the two has often been noted (e.g., Dixon 2012), but the best analysis of this relation has been given by Diessel (2003). Consider some examples from the Munda (Austroasiatic) language Kharia spoken in eastern and central India (Peterson 2011: 178-179, 183-184). Demonstratives and interrogatives have parallels both in inflection (e.g. *a=te* ‘which=OBL’, *u=te* ‘this=OBL’), and derivation (e.g., *a=ti’j* ‘which=side’, *u=ti’j* ‘this side’). Languages differ from each other in how strongly developed they are and how many interrogatives and demonstratives take part in the parallel development. Kharia, for example, has yet another interrogative (e.g., *i=te* ‘what=OBL’) as well as two (and formerly three) additional demonstratives (e.g., *ho=te* ‘that.MED=OBL’, *han=te* ~ *hin=te* ‘that=OBL’), not counting a loan from a neighboring language. Diessel (2003: 635) has shown that demonstratives, like interrogatives, “cross-cut the boundaries of several word classes”, express basic semantic categories (e.g., Kharia *ti’j* ‘side’ etc.), have etymologically non-analyzable stems, are not derived from but reinforced by lexical items, and share a similar pragmatic function (Diessel 2003; 2006): “both types of expressions are commonly used as directives that **instruct the hearer to search** for a specific piece of information outside of discourse (i.e. in the surrounding situation or in the hearer’s knowledge store).” (Diessel 2003: 636, my boldface) One difference between the two elements seems to be the fact that, while demonstratives are usually accompanied by a pointing gesture (Diessel 2006), this does not appear to be the case for most interrogatives. Although there are deictic interrogatives, they have a schematic meaning that contradicts a specific pointing gesture. In German discourse, however, in some cases a

selective interrogative can be accompanied with a pointing gesture, but this usually goes along with looking at the addressee and furling one's eyebrows or similar indicators of doubt. Whether there are more specific connections between interrogatives and gestures remains to be investigated.

### 4.4 Towards an ecological theory of questions

One of the questions formulated in the Introduction (Chapter 1) concerned the actual meaning of questions (Sanitt 2011: 561). Inspired by Schulze (2007) and van der Auwera & Nuyts (2007), this section thus goes beyond traditional typology and explicitly tries to add several theoretical assumptions concerning the *meaning* of questions and sketches what might be called an ecological theory of questions.

As noted in the Introduction, the fundamental unit for an ecological theory of language forms the so-called *organism-environment system* (OES, Järvillehto 1998). Many cognitive approaches overemphasize the importance of the organism and especially the brain. As Ulric Neisser—the so-called father of Cognitive Psychology—said in an interview in 1997, his 1976 book “*Cognition and Reality* was partly an attempt to recall my information processing colleagues to reality, saying that there is a whole world out there to look at.” (Szokolszky 2013: 187) However, Neisser also correctly pointed out that traditional Ecological Psychology (e.g., Gibson 1979) overemphasized the environmental aspect, but neglected memory and conceptualization. The theory of the organism-environment system, in my opinion, should aim at integrating aspects of both fields. The OES exists on several different time scales or causal frames (Enfield 2014) and contains language as an integral component (e.g., Odling-Smee & Laland 2009; Sinha 2013). However, in the remainder of this section a focus will lie on the understudied microgenetic frame. Some results from the diachronic and synchronic perspectives will be taken as hints of the basic infrastructure of this frame. This should not lead to the misunderstanding, however, that basic elements of the *human interaction engine* (Levinson 2006) or the *economics* of questions (Levinson 2012a), most of which are located on the enchronic frame and in the sociocultural ecology, are unimportant. This section merely focuses on some of the less well understood aspects of questions and emphasizes the microgenetic frame and the cognitive ecology of language (Steffensen & Fill 2014: 7). Graesser (1985: 3) was probably right that “a theory of questioning is a special case of a more general theory of conversation”, which is why only some aspects can be addressed here. Given the background of this book, this section is written from a linguistic perspective, although insights from other disciplines are consulted whenever feasible (cf. Dillon 1982).

Despite its ecological background, the general outline of the theory advocated here nevertheless is strongly based on the newly emerging *simulation semantics* paradigm that places a focus on the brain, but can easily be reconciliated with ecological ideas. The fundamental concept of this theory is so-called *embodied simulation*, which has been defined as “the re-enactment of perceptual, motor and introspective states acquired during experience with the world, body and mind” by Barsalou (2009: 1281) or as “the creation of mental experiences of perception and action in the absence of their external mani-

festation” by Bergen (2012: 14). These two definitions are more or less congruent and highlight different aspects of one and the same phenomenon. A definition offered by Gallese (2009: 527) in addition emphasizes the social aspect of simulations:

By means of embodied simulation we do not just “see” an action, an emotion, or a sensation. Side by side with the sensory description of the observed social stimuli, internal representations of the body states associated with these actions, emotions, and sensations are evoked in the observer, “as if” he or she were doing a similar action or experiencing a similar emotion or sensation. That enables our social identification with others.

Given its neurological background, the theory may be misunderstood as focusing on the brain, exclusively. However, Barsalou (2009) has emphasized that simulations are always situated and multi-modal, which is in accordance with the theory of the OES. The theory is broad enough to bring together conception, perception, and action (and thus the organism and the environment) into one coherent theory. According to Barsalou (2009: 1281)

the re-enactment process has two phases: (i) storage in long-term memory of multi-modal states that arise across the brain’s systems for perception, action and introspection (where ‘introspection’ refers to internal states that include affect, motivation, intentions, metacognition, etc.), and (ii) partial re-enactment of these multi-modal states for later representational use, including prediction.

Thus, simulations are never complete re-enactments but are *attenuated* to different degrees (Langacker 2008: 536-537).

It is especially the last aspect of a *prediction* or an *anticipation* (Järvillehto 2009) that plays a crucial role for a theory of questions. Every question (rhetorical questions etc. aside) contain aspects that are not actually known by the speaker but merely predicted or anticipated to play a role within a certain context. Assuming the hearer is cooperative (Tomasello 2014b), the question may be answered or responded to in an expected way, if the anticipation turns out to be appropriate. For example, one of two specified alternatives of an alternative question (45a) may be chosen as adequate and thus (partly) repeated by the hearer (45a). If, however, the anticipation was inadequate, then the hearer will most likely point this out and give the appropriate alternative (45c) or try to find out what the misunderstanding is about (45d).

(45) English

- a. *When are you leaving, tomorrow or the day after tomorrow?*
- b. *(I’m leaving) tomorrow.*
- c. *I’m not leaving, it is Bill who is leaving.*
- d. *I’m not leaving at all, what are you talking about?*

This is traditionally known as *presupposition* of a question. The background of these predictions has been called the *pattern completion inference mechanism*.

#### 4 The typology of questions

On encountering a familiar situation, an entrenched situated conceptualization for the situation becomes active. Typically, though only part of the situation is perceived initially. A relevant person, setting, event or introspection may be perceived, which then predicts that a particular situation—represented by a situated conceptualization—is about to unfold. By running the situated conceptualization as a simulation, the perceiver anticipates what will happen next, thereby performing effectively in the situation. The agent draws inferences from the simulation that go beyond the information given (Barsalou 2009: 1284)

Polar, focus, and alternative questions all rely on this anticipatory mechanism. The difference among them has to do with the fact that predictions may be more or less plausible, with the consequence that the information given may lead to one or more possible outcomes. In addition, the uncertainty may only concern a certain subpart of the entire simulation. This is one aspect of what is usually referred to as *construal* (e.g., Langacker 2008), the ability to “construe the ‘same’ situation quite differently” (Ross 2014 [1987]: 127). Content questions lack any specific predictions but still involve inferences in the sense that they rely on the activation of entrenched situated conceptualization. Consider the example of a broken window. We know from our previous experience that windows usually don’t break on their own and that somebody or something must have caused the glass to break. Most likely we would assume that there must be an agent responsible for breaking the window (e.g., one of the children usually playing soccer in front of the house), leading to the question *Who broke the window?* In case we have encountered a similar situation before and know the identity of a potential agent, we may also ask something like *Did Tom break the window again?* Questions are an expression of the human imaginative capacity and thus, ironically, of knowledge, memory, and experience.

Tomasello (2008: 84–87) differentiates between three basic communicative motives, i.e. *requesting*, *informing*, and *sharing*. Arguably, questions can be used for all three motives. Consider the constructed examples in (46).

(46) English

- a. *Could you open the window?*
- b. *Did you know Sarah is pregnant?*
- c. *That’s beautiful, isn’t it?*

Given the overall focus of this study, however, only prototypical questions can be covered here, i.e. actual requests for information (e.g., Levinson 2012a), which is a special case of the first motive. However, as we have just seen, every question itself necessarily contains some amount of information.

Have you ever hesitated to ask a question? Perhaps you feared it might be foolish. Or it might be too near the bone, too probing. Perhaps it might cause offence. Or it might distract us from the business at hand and lead to other things. Or it might open you up to the reciprocal question, which you would not want to answer. Introspection suggests a plethora of reasons for suppressing questions that might arise in one’s mind. (Levinson 2012a: 19)

In a certain sense, questions are an example of the *perception-action cycle* as postulated in Ecological Psychology: “animals [including humans] move so that they can perceive, and perceive so that they can move” (Swenson & Turvey 1991: 319, my brackets). Questions give information in order to obtain additional information necessary for a certain purpose. Nevertheless, prototypical answers are a better example of the second communicative motive. Interestingly, requesting appears to precede informing both phylogenetically and ontogenetically (Tomasello 2008: 137, 247) and thus clearly plays a fundamental role for human beings. The third motive is irrelevant for the purpose of this study.

Prototypical questions may furthermore be characterized as a form of *exploratory behavior* that results from *curiosity*. According to Loewenstein (1994: 87), curiosity in the sense of “an intrinsically motivated desire for specific information” is raised by the focusing of a gap in our knowledge base. Such “an information gap is characterized by two quantities: what one knows and what one wants to know.” All question types may be characterized in the same terms. In content questions the entrenched situated conceptualization equips us with a schematic knowledge but inquires about a specific piece of information one wants to know. In the case of *who*, we know about an agent but want to know its identity. In polar and focus questions we have a specific assumption but do not know whether this is accurate. In alternative questions we can imagine two or more possibilities but do not know which one is the most accurate. The underlying pattern can be called a *hierarchy of specificity* of question types (47, cf. Levinson 2012a: 23; Hölzl 2016a).

(47) CQ < PQ < FQ < AQ

The term *specificity*, which contrasts with *schematicity*, has been adopted from Langacker (2008: 19); see also Arnheim (1969: 238). Focus questions are more specific than polar questions, because the uncertainty just concerns the focused subpart which is much more specific than in content questions. Alternative questions appear to be the most specific, because they openly specify all plausible alternatives. The possible negative answer in polar and focus questions opens up a plentitude of alternatives. There is direct evidence for this hierarchy. One pattern recurring in many languages is a combination of a content question followed by a polar, focus, or alternative question that elaborates on the frame set by the content question (e.g., *What do you want, coffee or tea?*). Consider the following examples from Northeast Asia (48–50) and beyond (51–53).

(48) Evenki (Tungusic)

*si i:-le ɲene-d'e-nni, [d'u-la-vi=gu, tatkit-tula=gu]?*  
2SG which-ALL go-PRS-2SG home-ALL-REFL.POSS=Q school-ALL=Q

‘Where are you going, [home or to school]?’ (Nedjalkov 1997: 7)

(49) Khorchin Mongolian (Mongolic)

*čii jaa.x-sə=ji, [tolgɔ=čɪn' ubud-jææ-n=vs]?*  
2SG do.what-P.PFV=Q head=2SG.POSS hurt-PROG-PRS=Q

‘What’s up, [is your head aching]?’ (Yamakoshi 2015: 296)

#### 4 The typology of questions

- (50) Udihe (Tungusic)  
*i:-le una-za-i [amä:-za-la=nu zylie-ze-le=nu]?*  
 which-LOC travel-SUBJ-2SG back-N-LOC=Q front-N-LOC=Q  
 ‘Where will you travel, [in the front or in the back]?’ (Nikolaeva & Tolskaya 2001: 812)
- (51) Teiwa (Timor-Alor-Pantar)  
*amidan la qau, [tii’ le karian]?*  
 what FOC good sleep or work  
 ‘What is better, [sleeping or working]?’ (Klamer 2010: 284)
- (52) Amis (Austronesian)  
*cima ku ta.tayra namu i taypak, [ci luŋi ci áki]?*  
 who NOM go 2PL.GEN PREP PN NOM PN NOM PN  
 ‘Which one of you is going to Taipei, [Lungi or Aki]?’ (Huang et al. 1999: 650)
- (53) Central Alaskan Yupik (Eskaleut)  
*nali-ak assik-siu, [kuuvviaq wall’u saayuq]?*  
 which-ABS.3DU.SG like-2SG.3SG.Q coffee.ABS.SG or tea.ABS.SG  
 ‘Which do you want, [coffee or tea]?’ (Miyaoaka 2012: 170)

It is difficult to determine whether this is a universal pattern, because grammar books never explicitly address it as a phenomenon on its own right. Nevertheless, the fact that it can be found in languages from around the world indicates that it is a strong tendency at the very least. Future studies have to determine the exact meaning of this pattern, which may differ from instance to instance and from language to language. Additional examples from Chalkan, Chuvash, Udihe, Uilta, Uzbek, Kalmyk, and Ket can be found throughout Chapter 5. See also §6.3 for examples from the Timor-Alor-Pantar language Abui and the Austronesian language Balantak. In general terms the pattern can be described as the iconic linguistic expression of a possible universal that starts with the schematic and, by means of exploration and anticipation, gradually arrives at the more specific (e.g., Bar 2009; Barsalou 2009). The same phenomenon can be observed in focus questions with a focus on generic nouns that are more specific than interrogatives, but are followed by a question with an even more specific or proper noun (e.g., *Do you want tea, Earl Grey or Pu-Erh perhaps?*). In both cases the crucial point is that the first question is located lower on the scale of specificity in (47) than the second. In a way, epistemic tag questions mirror this structure because they start from a rather general statement and arrive at the specific question of whether this statement is appropriate (e.g., *You want tea, right?*). A major difference, however, is the fact that the first element in a tag question is not a question itself or at least is not overtly marked as such. Another difference is the scope of the second question over the whole proposition in the case of many tag questions. Alternative questions exhibit some affinity to this pattern as well, but there are major differences. While in all examples above the second sentences elaborate on, or are based on, the first one, alternative questions have mutually exclusive alternatives. A similarity

with alternative questions is, however, the fact that in both cases there is the possibility of ellipsis (cf. *Do you want tea or (do you want) coffee?* and *What do you want, (do you want) coffee or tea?*). Nevertheless, alternative questions are better treated as a question type comparable to polar and content questions as there are further differences, such as the connection of alternative questions with the domain of coordination. Tag questions, which are much more similar to this pattern than alternative questions, of course, do not repeat the same statement. Instead, question tags may anaphorically refer to the statement (e.g., *isn't it?*).

To borrow a term from Langacker (2008) again, it may be claimed that these combinations of questions follow a natural and dynamic path of *mental access* that unfolds through time, from the schematic to the specific:

Between the moment the organism is confronted with the problem and the moment the final solution is achieved there occur, as a rule, a number of intermediate steps leading, in an hierarchical fashion, **from general to more specific features** of the sought-after solution. (Duncker & Krechevsky 1939: 178, emphasis modified)

In Langacker's (2008: 83) terminology, this can also be called a *reference point relationship*, in which the second part (the target) is mentally located with respect to the first (the reference point). Dewey's (1910: 102) description of the phenomenon is still surprisingly accurate. He differentiates between three different situations, the first two of which define the extremes, i.e. absolute certainty and uncertainty:

Unless there is something doubtful, the situation is read off at a glance; it is taken in on sight, *i.e.* there is merely apprehension, perception, recognition, not judgment. If the matter is wholly doubtful, if it is dark and obscure throughout, there is a blind mystery and again no judgment occurs.

The third situation exactly corresponds to the scale of uncertainty in between these extremes:

But if it suggests, however vaguely, different meanings, rival possible interpretations, there is some *point at issue*, some *matter at stake*. Doubt takes the form of dispute, controversy; different sides compete for a conclusion in their favor. Cases brought to trial before a judge illustrate neatly and unambiguously this strife of alternative interpretations; but any case of trying to clear up intellectually a doubtful situation exemplifies the same traits. A moving blur catches our eye in the distance; we ask ourselves: "What is it? Is it a cloud of whirling dust? a tree waving its branches? a man signaling to us?" Something in the total situation suggests each of these possible meanings. Only one of them can possibly be sound; perhaps none of them is appropriate; yet *some* meaning the thing in question surely has.

Not only this combination of questions, but questions in general can be characterized as an expression of *uncertainty* (e.g., Schulze 2007). However, uncertainty is merely one of several *collative variables*, a term coined by Berlyne (1960: 44).

#### 4 The typology of questions

For want of a more satisfactory term, we shall call them *collative variables* since, in order to evaluate them, it is necessary to examine the similarities and differences, compatibilities and incompatibilities between elements—between a present stimulus and stimuli that have been experienced previously (novelty and change), between one element of a pattern and other elements that accompany it (complexity), between simultaneously aroused responses (conflict), between stimuli and expectations (surprisingness), or between simultaneously aroused expectations (uncertainty).

Given the ecological background of this study, the terms *stimulus* and *response* have to be treated with caution. Instead of passively reacting to the environment, the organism itself may engage in active exploratory behavior (e.g., Dewey 1896; 1910: 193; Gibson 1960; 1979: 55ff.; Gibson 1988: 5-6). Conceptually, this is a similar distinction as that between *natural selection* by the environment and *niche construction* by the organism that we have seen in the Introduction (Odling-Smee & Laland 2009). In many cases, it is the actions and the movements of the organism itself that lead to the pick-up of *novel, changing, complex, conflicting, surprising, or uncertain* information. Baranesa et al. (2015: 89) argue “that **curiosity** can be viewed as a pro-active process that anticipates, or motivates agents to obtain new information, whereas **surprise** indicates a reactive process after having processed the information.” (my boldface) This in turn results in further exploratory behavior.

Of course, there is also the artificial arousal of curiosity such as, for instance, in a *riddle*, which “compares an object to another entirely different object. Its essence consists in the surprise that the solution occasions”. Eventually, “the hearer perceives that he has entirely misunderstood what has been said to him.” (Taylor 1943: 129) The riddle arouses curiosity in the addressee by means of collative information and initiates the search for the solution. Take an example from the Tungusic language Uilta, which starts with the introduction *gaŋ gaŋ gajagoo!* and goes on as follows: *boo toptongoor, naa toptongoor, xai-gæək? toksiik unuu!* ‘In heaven there are spots, on earth there are spots, what are they? Riddle me!’ The riddle has several possible answers such as *boo unigərinnii suŋdatta xəsiktənnii* ‘The stars in heaven and the scales of fish.’ The answer is followed by the reply *toksiik* ‘Correct.’ (Ikegami 1958: 93), which puts an end to curiosity.

A basic typology of different kinds of curiosity was also sketched by Berlyne (1954) who differentiates between two dimensions that define four types of curiosity (see also Dewey 1910: 30ff.). These have been concisely summarized by Loewenstein (1994: 77) as follows.

**Perceptual** curiosity referred to “a drive which is aroused by novel stimuli and reduced by continued exposure to these stimuli” [Berlyne 1954: 180]. **Epistemic** curiosity referred to a desire for knowledge and applied mainly to humans. **Specific** curiosity referred to the desire for a particular piece of information, as epitomized by the attempt to solve a puzzle. Finally, **diversive** curiosity referred to a more general seeking of stimulation that is closely related to boredom. In the four-way categorization produced by these two dimensions, **specific perceptual** curiosity is

exemplified by a monkey's effort to solve a puzzle, **diversive perceptual** curiosity is exemplified by a rat's exploration of a maze [...], **specific epistemic** curiosity is exemplified by a scientist's search for the solution to a problem, and **diversive epistemic** curiosity is exemplified by a bored teenager's flipping among television channels. (my boldface and square brackets)

It is especially *specific epistemic curiosity* that plays a crucial role for the characterization of questions. Above, we have already encountered the knowledge gap theory of curiosity by Loewenstein (1994), which is strongly based on Gestalt Psychology: "If curiosity is like a hunger for knowledge, then a small 'priming dose' of information increases the hunger, and the decrease in curiosity from knowing a lot is like being satiated by information." (Kang et al. 2009: 963) The first to sketch a gestalt approach to curiosity was also Berlyne (1954: 181), proposing "a drive to fill in such gaps in the subject's experienced representations". This is based on the well-known gestalt principle of *closure*. Fritz Perls (1973: 119)—the father of Gestalt Therapy—put it this way: "The gestalt wants to be completed. If the gestalt is not completed, we are left with unfinished situations, and these unfinished situations press and press and press and want to be completed." In a different terminology one could say that an embodied simulation wants to be completed. For instance, unanswered questions usually lead to an "increased effort in constructing a coherent representation" (Hoeks et al. 2013: 8). If there is insufficient information to complete a simulation, curiosity and exploration set in. What exactly the evolutionary origins of curiosity are is another matter that cannot be addressed here. The point is that curiosity is a psychologically real phenomenon and has to be taken into account for a characterization of questions. Gibson's (1979: 219) statement that "[t]he visual system *hunts* for comprehension and clarity" can perhaps be generalized to the entire organism-environment system. Humans seek comprehension and clarity, and questions are one way of achieving this. Berlyne's (1954: 182) description is based on a somewhat outdated terminology but nevertheless remains basically valid:

When a question is put, whether by the subject himself or by somebody else, and the answer is already known, the appropriate response is made as a reaction conditioned by previous learning to the stimulus-pattern, and this relieves the drive immediately, so that the subject can proceed to some other activity. However, when the answer is not known, the drive will persist, and some sort of trial-and-error process can be expected to follow as with any other drive-state.

He mentions three different possibilities for this "trial-and-error process", *thinking, observation, and recourse to authority*. The first refers to processes mostly restricted to the organism such as problem solving or memory, but the latter two roughly correspond to the physical and social environment, respectively (see also Lewin 1936: 24ff.; Steffensen & Fill 2014: 7).

Put differently, one may *resolve curiosity* in three different but interrelated ways. First, in most cases one's own experience and memory are sufficient, although in some cases additional thought processes such as problem solving may be necessary. This is the tra-

#### 4 *The typology of questions*

ditional realm of Cognitive Science. In his *Natural history of human thinking*, Tomasello (2014a) defines *thinking* as

a single cognitive process, but one that involves several key components, especially (1) the ability to cognitively represent experiences to oneself “offline”; (2) the ability to simulate or make inferences transforming these representations causally, intentionally, and/or logically; and (3) the ability to self-monitor and evaluate how these simulated experiences might lead to specific behavioral outcomes—and so to make a thoughtful behavioral decision.

The fundamental mechanism of representation assumed in this study is embodied simulation as defined above. Perhaps most instances of curiosity are simply resolved by inferences and predictions and their subsequent evaluation whether they are plausible or not. But it is wrong to assume, as Tomasello seems to be well aware, that simulations may be completely “offline” or detached from the environment. In fact, as Glenberg (1997: 1) observed, simulations may be said to be basically “driven by the environment”:

A significant human skill is learning how to suppress the overriding contribution of the environment to conceptualization, thereby allowing memory to guide conceptualization. The effort used in suppressing input from the environment pays off by allowing prediction, recollective memory, and language comprehension.

The pay-off is a plausible evolutionary explanation to pay less attention to a potentially dangerous environment. But Glenberg’s (1997) inclusion of language comprehension is problematic, as language usually is an aspect of our social environment. Simulations based on language, for example when we listen to somebody asking us a question, are certainly driven by the (social) environment.

Second, in some cases we may encounter situations or objects that we have not encountered before or are otherwise unfamiliar with. In this case we may simply move around, explore, and change our relative perspective and distance in order to perceive previously inaccessible aspects. This is something Ecological Psychology has focused on from its very beginnings (see Gibson 1979). In this sense, curiosity is simply resolved by physically exploring and changing our position relative to the problematic object. Small objects, of course, may be grasped and turned in order to be investigated in a more thorough fashion. A different example of physical exploration based on diversive instead of specific curiosity can be illustrated by the wanderlust of the Tungusic speaking Evenki in Siberia.

The Evenki learn from an early age **to be interested in, rather than frightened by, risky situations** and the possibility of exploring new territories. For them, seeking out new places offers a wonderful opportunity to experience companionship and, as a result, it is common to go somewhere **just for the sake of exploration**. (Safonova & Sántha 2013: 142, my boldface)

This is one of several reasons for the extraordinary wide distribution of Evenki and their close linguistic relatives such as the Even over all of the northern half of NEA (§2.10).

Third, instead of observing or thinking on our own, we may also see whether other people can help us clarify certain aspects of a problematic object or situation. This is mostly accomplished by means of language, of course, and the most prototypical tool for this are questions. Hodges (2009: 636), in analogy to Gibson's (1979) *ambient optic array*, proposed the name *dialogical array*,

a group of hearer-speakers surrounding a given speaker-hearer, listening and talking in ways that reveal, inevitably, something of their perspectives, their intentions, and their histories relative to the present place and time. Like light, the ordered gestures of the array, as well as their disordering and reordering over time, allow a participant in the array to have their own orderings restructured on various scales. It is an array of partners, actual and potential, who provide information, not just about themselves as intentional agents and as objects, but about objects, events, and agencies beyond the physical and temporal horizons of the immediate physical surround. (Hodges 2009: 636)

The dialogical array *affords* (Lewin 1936; Gibson 1979) linguistic interaction such as asking questions. From one point of view questions are a form of bodily action that bring about changes in the dialogical array, which in turn allows the pick-up of new information (cf. Swenson & Turvey 1991). This reliance on other people potentially brings with it the danger of deception and misinformation as well as of social costs (e.g., Levinson 2012a: 20), but pays off by being faster and requiring less effort, especially if we are dealing with complex problems. This might be the reason why questions apparently are a universal property of language. While exploratory behavior can also be found in other animals, language in general and questions in particular crucially depend on the *ultra-social* nature of human beings who usually tend to cooperate with each other in ways that are unique (Tomasello 2014b).

Of course, the above distinction is only a heuristic one. In principle, the three means of resolving curiosity are interrelated and often combined. They merely highlight different aspects of the organism-environment system (Lewin 1936: 27; Steffensen & Fill 2014: 7). Questions, for instance, necessarily contain aspects of all three types of exploration mentioned above. While the social dimension is the most important, both physical movements (e.g., eye contact) as well as thinking (e.g., predictions) are crucial elements as well. Questions trigger incomplete simulations in the hearer, based on her experience and memory, who then engages in exploratory behavior herself. Here basically the same three mechanisms come into play. Either the hearer has sufficient information to fill in the gaps herself, or she engages in other exploratory behavior (e.g., looking something up), or seeks additional help and asks the same question of somebody else who is likely to know the answer.

The discussion thus far has overemphasized the microgenetic aspect of questions, but this last point has mentioned some enchronic properties as well. The social aspect of questions can be observed, for example, through a relatively strong obligation on the part of the addressee to respond (e.g., Levinson 2012a: 16). The interaction of questions with evidentiality offers additional insights into the social nature of interrogativity. In

#### 4 The typology of questions

many languages that have perspective marking, there is a shift to the perspective of the addressee in questions. Consider some examples from the Cha'palaa language spoken in Ecuador that has an egophoric system.

(54) Cha'palaa (Barbacoan)

a. (*i-ya*) *pipe-yu*.

1-TOP bathe-EGO

'I bathed.'

b. (*ñu/ya*) *pipe-we*.

2/3 bathe-N.EGO

'You/(s)he bathed.'

c. (*ñu-ya*) *pipe-yu?*

2-FOC bathe-EGO

'Did you bathe?'

d. (*ya-a*) *pipe-n?*

3-FOC bathe-N.EGO.Q

'Did (s)he bathe?'' (San Roque et al. 2017: 136-137)

The egophoric marker *-yu* appears in both statements that refer to a first person (54a) and in questions that refer to a second person (54c). Tournadre & LaPolla (2014: 245) capture this phenomenon with the *anticipation rule*, which they illustrate with Tibetan: “when ever the speaker asks a direct question of the hearer, she should anticipate the access/source available to the hearer and select the evidential auxiliary/copula accordingly.” The underlying mechanism can be explained with the help of embodied simulation: The questioner asks the question *as if* she was the addressee herself (Gallese 2009: 527), using predictions obtained through mentally simulating the situation. See §5.9.2.1 on Wutun and §5.9.2.2 on Amdo Tibetan for additional examples.

According to Schulze (2007: 248), furthermore, there is a “strong coupling of the first person with assertions and of the second person with modal features, among them interrogativity.” This important observation, it seems, can be directly observed in a number of languages that exhibit a split type based on person. Qiang, for example, which we have encountered above, has a special question marker for second person singular. Some Turkic languages have a split system that is sensitive to second person, too (§5.11.2). Regarding West Greenlandic (Eskaleut), Sadock (1984: 199) observed that

in all cases where the subject is second person, there is an interrogative form that is distinct from the indicative; in some cases where the subject is third person (nowadays only where there is no object, but formerly also where the object was third person), there are distinct interrogative and indicative forms; but in no case where the subject is first person is there a separate interrogative form.

Sadock proposes the hierarchy in (55).

(55) 2 > 3 > 1

Questions are most likely to refer to a second person and least likely to refer to a first person. This should not be interpreted as a strict implicational hierarchy, however, which allows no exceptions. Nevertheless, it seems to be a valid tendency that had actually already been discovered by Bolinger (1957: 3): “*You* occurs oftener than not in Qs. It therefore ‘means’ ‘question,’ loosely and insufficiently, but enough so that a locution not otherwise identifiable as a Q becomes one (is reacted to as one) if *you* is present.” This highlights the social aspect of questions, which are strongly rooted in communicative interaction, and has an analogue in the gazing behavior of the questioner. Rossano et al. (2009: 239), based on the investigation of the three very different speech communities of Italian (Indo-European), Yéli Dnye (no affiliation), and Tenejapan Tzeltal (Mayan), found that it is especially the questioner who is gazing at the addressee (instead of the other way around), which is in accordance with my subjective impression for conversations in German. Recently, Baranesa et al. (2015: 81) additionally found “that higher curiosity was associated with earlier anticipatory orienting of gaze toward the [expected] answer location”. These facts are also consistent with the explanation of questions as a form of exploratory behavior in the dialogical array because most other types of exploration involve some kind of active looking. As Gibson (1979: 212) put it, “looking is always exploring”.



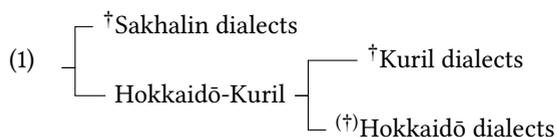
# 5 Survey of the grammars of questions in Northeast Asia

This chapter takes a closer look at the grammar of questions in language families of NEA in alphabetical order, from Ainuic (§5.1) to Yeniseic (§5.14). Each section on a language family is divided into three parts, a brief introduction that sketches its internal classification, a section on question marking, and one on interrogatives. For practical purposes, subsections in the larger language families Indo-European (§5.5) and Trans-Himalayan (§5.9) are distinguished additionally into subbranches such as Germanic or Sinitic. The part on Yeniseic has an additional subsection on the Dene-Yeniseian hypothesis (§5.13.4). Please note that, except for perhaps Tungusic, the classification of each language family is not exhaustive and is mostly intended as a tool that allows to better understand the internal order of the individual subsections.

## 5.1 Ainuic

### 5.1.1 Classification of Ainuic

Ainuic has three dialect groups that are named after their geographical distribution. These are the Sakhalin dialects, Kuril dialects, and Hokkaidō dialects. Excluding the possible existence of now extinct Para-Ainuic varieties on Honshū, the Ainuic language family may roughly be classified as follows (cf. Vovin 1993: 157, see also Figure 5.1 in §5.1.3).



According to Shibatani (1990: 4) there is what he calls “Classical Ainu”, the language of oral epics (*yukar*), which differs from the spoken language and allegedly represents older stages of development. But Nakagawa & Okuda (2007: 378) claim that

it is misleading to describe the grammar of Ainu as resting upon this distinction, because the behaviour and distribution of so-called “classical” features are actually independent from each other. There is no sound evidence to support the claim that the “classical” features are really older than “colloquial” ones in the history of this language.

The following description of question marking is mostly based on the Saru and Chitose dialects in southwestern Hokkaidō as well as the Shizunai and Tokachi dialects in southeastern Hokkaidō. The Sakhalin and Kuril dialects will only be mentioned briefly. A more complex picture including almost all dialects can be drawn for the interrogative system.

### 5.1.2 Question marking in Ainuic

For marking polar questions, the Saru dialect of Ainu has final rising intonation combined with an optional final particle *ya*.

- (2) Ainu (Saru)  
*nisatta nupurpet or un e=arpa ya?*  
 tomorrow PN place ALL 2SG.S=go.PL Q  
 ‘Will you go to Noboribetsu tomorrow?’ (Bugueva 2012: 497)

The online *Topical Dictionary of Conversational Ainu* based on the Saru dialect contains a section called *Question and Answer* from which the following example of an unmarked polar question with a longish and slightly rising intonation towards the end was drawn.

- (3) Ainu (Saru)  
*ku=ye itak e=raman?*  
 1SG.A=say language 2SG.A=know  
 ‘Do you understand what I am saying?’ (NINJAL 2015)

Interrogatives are *in situ* and there is usually no additional morphosyntactic marking, though the use of *ya* is possible. Alternative questions exhibit double marking with *ya*.

- (4) Ainu (Saru)
- a. *hunna ek?*  
 who come  
 ‘Who came?’
  - b. *hemanta e=e rusuy ya?*  
 what 2SG.S=COP want Q  
 ‘What would you like to eat?’ (Tamura 2000: 235, 236)
  - c. *ek ya, somo ya?*  
 come Q NEG Q  
 ‘Are (you) coming or not?’ (NINJAL 2015)

The use of the same final particle for both polar and content questions has an areal connection to surrounding languages (§6). Alternative and (truncated) polar questions are both marked with *=he*.

- (5) Ainu (Saru)  
*matci=he e=kor rusuy tampaku=he e=kor rusuy?*  
 match=Q 2SG.S=have want cigarette=Q 2SG.S=have want  
 ‘Do you want the matches or the cigarettes?’ (Tamura 2000: 234)
- (6) a. *toan pe ekte.*  
 that thing hand.over  
 ‘Give that to me.’  
 b. *tan pe=he?*  
 this thing=Q  
 ‘This one?’ (Tamura 2000: 234)

In one recorded open alternative question, perhaps because =*he* cannot combine with interrogatives, only the first alternative takes the marker =*he*.

- (7) Ainu (Saru)  
*te ta=he, hunak ta?*  
 here LOC=Q where LOC  
 ‘Here or where?’ (NINJAL 2015)

An additional marking of polar questions grammaticalized from nominalization has parallels in Japanese (see §5.6.2). According to Bugaeva (2012: 497) the final copula *ne* may be omitted following the evidential infinitive marker *ruwe* (*ne*) ‘it is a fact that’, which in turn seems to mark polar questions. The same pattern can be observed in the Chitose dialect (Bugaeva 2004: 85). Tamura (2000: 233) claims that the same development is also possible with other evidential markers, notably *hawe* (*ne*) ‘it is said that’ and *siri* (*ne*) ‘it looks that’. It appears that the newly grammaticalized question markers may be present in polar, alternative, and content questions.

- (8) Ainu (Saru)
- a. *yosiku e=ne ruwe?*  
 PN 2SG.S=COP Q  
 ‘Are you Yoshiko?’ (Bugaeva 2012: 497)
- b. *na tuyma ruwe, hanke ruwe?*  
 still far Q be.close Q  
 ‘Is it far or near?’
- c. *makanak pak-no sir-tuyma ruwe?*  
 what till-ADV appearance-be.far Q  
 ‘How far is it?’ (NINJAL 2015)

The translation of the three markers above was taken from Bugaeva’s (2012: 494) description, which contains yet another evidential marker *humi* (*ne*) ‘it feels that’ (see 17 and 24 below). The nominalizers or evidential markers transparently derive from nouns, namely

## 5 Survey of the grammars of questions in Northeast Asia

“inferential *ruw-e* (< ‘the trace of’), reportative *haw-e* (< ‘the voice of’), non-visual (= semblative) *hum-i* (< ‘the sound of’), visual *sir-i* (< ‘the sight of’)” (Bugaeva 2012: 470). The evidential markers also appear in what seem to be tag questions, where they are followed by *somo ne (ya)*. The use of the question marker is optional.

- (9) Ainu (Saru)  
*e=sinki ruwe somo ne (ya)?*  
 2SG.S=be.tired INF.EV NEG COP Q  
 ‘You’re tired, aren’t you?’ (Tamura 2000: 233)

In addition to the question markers mentioned above, there is a special copula *an* that replaces the plain copula *ne* in questions. Special interrogative copula forms are also known from several Mongolic languages (§5.8.2) as well as Shuri (§5.6.2).

- (10) Ainu (Saru)  
*núman hunna ek ruwe an?*  
 yesterday who come.SG INF.EV COP.Q  
 ‘Who came yesterday?’ (Bugaeva 2012: 497; Tamura 2000: 237)

For the **Chitose** dialect Bugaeva (2004: 88) mentions the fact that the special copula is usually encountered after one of the evidential markers mentioned above, though see example (17b) for a counterexample of the Tokachi variety. The copula can also appear twice in alternative questions.

- (11) Ainu (Saru)  
*ooho pet an, ohak pet an?*  
 be.deep river COP.Q be.shallow river COP.Q  
 ‘Is it a deep or a shallow river?’ (NINJAL 2015)

According to Batchelor (1905: 141), the enclitic *=he* “expresses interrogation, and is often though by no means always, followed by the verb *an* ‘to be.’”

- (12) Ainu (Saru)  
*tan kur aynu itak eraman kur=he an?*  
 this person PN language know person=Q COP.Q  
 ‘Does this person understand Ainu?’ (NINJAL 2015)

A question construction specialized for inquiring about topics is *hike (mak)?* ‘how about’ (Tamura 2000: 237), which appears to take a sentence-final position. As we will see in §5.1.3, *mak* is actually an interrogative meaning ‘how, why’, while *hike* is a conjunction with the meaning ‘and’ (Bugaeva 2012: 497).

In the **Chitose** dialect, few polar questions are marked with rising intonation alone. In most cases it is combined with the same final question marker *ya* as seen above.

- (13) Ainu (Chitose)  
*tan-to e=nepki humi pirka a ya?*  
 this-day 2SG.S=work EV.N be.good PFV.PL Q  
 ‘Did you work well today?’ (Bugueva 2004: 85)

Content questions in the Chitose dialect are also said to exhibit the marker *ya* more often, as compared with the Saru dialect (Bugueva 2004: 86). There are, nevertheless, content questions without the marker.

- (14) Ainu (Chitose)  
*eani hunna e=ko-ysoytak?*  
 2SG who 2SG.S=to.APPL-talk  
 ‘Who are you talking to?’ (Bugueva 2004: 86)

Alternative questions have the same particle *he* as seen in the Saru dialect above. But Bugueva mentions an example of an alternative question which in addition exhibits the question marker *ya* following each alternative. Altogether there are thus four question markers. Apparently, *=he* attaches to the focus, while *ya* can be found in final position after each alternative.

- (15) Ainu (Chitose)  
*seta=he ne ya, kamuy=he ne ya?*  
 dog=Q COP Q god=Q COP Q  
 ‘Is it a dog or a god?’ (Bugueva 2004: 88)

The **Shizunai** dialect also has the question marker *ya* in sentence-final position, which seems to have the same semantic scope as seen before.

- (16) Ainu (Shizunai)  
*numan ekasi nep kar ya?*  
 yestderday old.man what do Q  
 ‘What did the old man do yesterday?’ (Refsing 1986: 229)

As in the Saru and Chitose dialects, there is a connection of questions to nominalizations, i.e. *ruwe*, *siri*, *hawe*, and *pe*. The first three correspond to the Saru forms mentioned above while the last one is similarly neutral like *ruwe*. The difference between the two is the level of abstractness, *pe* referring to concrete and *ruwe* to abstract objects (e.g., Refsing 1986: 229f.). The copula *an* is attested as well.

A recent treatment of the **Tokachi** dialect in southeastern Hokkaidō mentions several questions that exhibit no significant difference from the other dialects already mentioned.

5 Survey of the grammars of questions in Northeast Asia

(17) Ainu (Tokachi)

- a. *nep tap eci-ki sir okay ya?*  
 what EMPH 2PL.A=do EV.N COP Q  
 ‘What are you doing?’
- b. *oupeka ekaci e=cakoko a ru etap an?*  
 straight child 2SG.A=teach PFV.PL EV.N EMPH COP.Q  
 ‘Did you teach the children correctly?’
- c. *e=cikir-i arka ru=he?*  
 2SG.S=foot-POSS hurt EV.N=Q  
 ‘Are your feet hurting?’ (Takahashi 2013: 131)

Based on this similarity, one may speculate that alternative questions presumably display double marking with =*he* and that the question marker *ya* also marks polar questions. Tokachi Ainu has yet another question marker *a* not encountered thus far. In all examples given, it follows the copula *an* and marks content questions.

(18) Ainu (Tokachi)

- nen tap apusta kik hum an a?*  
 who EMPH door knock EV.N COP.Q Q  
 ‘Who knocked on the door?’ (Takahashi 2013: 131)

Apparently, the marker also exists in other dialects such as Saru. The following example illustrates that it can also appear in polar questions.

(19) Ainu (Saru)

- arki rok a?*  
 come.PL PFV.PL Q  
 ‘Have they come?’ (Shibatani 1990: 79)

Other Hokkaidō dialects seem to exhibit a pattern very similar to those already observed, though there usually is only little information available. For example, the Samani dialect also has the marker *ya* and the special interrogative copula ‘*an*’, but additional information on further question markers and their semantic scope remain obscure (T. Tomomi 2002: 101, 107).

For Sakhalin Ainu, the materials collected by Konada (Tittel 1922) contain the three question markers *a*, *ya*, and *he*. We have already encountered all three markers above in several Hokkaidō dialects. Their semantic scope remains unclear but may be similar to Hokkaidō dialects as well.

(20) Ainu (Sakhalin)

- a. *pirika a?*  
 be.good Q  
 ‘Is it alright?’

- b. *tam-pe ne ya?*  
 this-thing COP Q  
 ‘Is it this thing?’
- c. *e=nu=he?*  
 2SG.A=hear=Q  
 ‘Do you hear?’ (Tittel 1922: 85)

For the **Kuril** dialect of Ainu, there is a content question that was originally recorded by Voznesenskii. Apart from the interrogative, no marking is present.

- (21) Ainu (Kuril)  
*nie-bie-gor?*  
 what-thing-have  
 ‘What thing is there?’ (Vovin 1993: 199)

No information on other question types in this dialect group seems to be available.

Table 5.1 summarizes the limited information of Ainuic question marking that we have seen above. The semantic differences between different markers of polar questions as well as the exact semantic scope for most forms remains obscure for now.

Table 5.1: Tentative summary of question marking in Ainuic

Language	PQ	CQ	AQ
Chitose Ainu	ya#	ya#	2x =he (+ 2x ya#)
Kuril Ainu	?	?-	?
Sakhalin Ainu	ya#, a# =he#	?	?
Saru Ainu	ya#, a# =he#, an COP.Q, N#	ya#, an COP.Q, N#	2x =he, 2x ya#, 2xN#
Shizunai Ainu	ya#, an COP.Q	ya#, an COP.Q	?
Tokachi Ainu	?ya#, an COP.Q	ya#, a#, an COP.Q	?

As usual, most question markers remain etymologically opaque, but Ainuic *ya* could be somehow related to Old Japanese =*ya* (§5.6.2). A problem for the comparison is, however, a different morphosyntactic behavior and semantic scope of the Old Japanese marker that is a mobile enclitic not found in content questions.

### 5.1.3 Interrogatives in Ainuic

The sets of interrogatives in the three dialects Saru, Chitose, and Shizunai mentioned in the previous section are very similar to each other (Table 5.2). For the Tokachi dialect Takahashi (2013) only mentions *nen* ‘who’, *nep* ‘what’, *nekon* ‘how’, and *onon* ‘whence’.

From a synchronic point of view, the interrogatives are mostly opaque, but at least some forms are readily analyzable. The form *nep kusu* ‘why’ from the Chitose dialect consists of *nep* ‘what’ and *kusu* ‘because’. The Shizunai dialect in this expression has

5 Survey of the grammars of questions in Northeast Asia

Table 5.2: Saru (Tamura 2000; NINJAL 2015), Chitose (Bugueva 2004), and Shizunai interrogatives (Refsing 1986)

Meaning	Saru	Chitose	Shizunai
who	hunna	hunna	hunna, nen
where (to)	hunak (ta), hinak (ta)	hunak	hunak, neyta
what	hemanta, hñta	hemanta ‘what, why’	hemanta
how much/many	hempak		hempak
when	hempara	hempara	hempara
why	hemanta ne	nep (kusu)	nepkus (ta)
which	inan, (h)inaan	inaan	
how, why	mak, makanak	neun, makanak	nekon

Table 5.3: Sakhalin Ainu interrogatives according to Bronisław Piłsudski (Majewicz 1998: passim) with tentative additional analysis based on Shibatani (1990) and Bugueva (2012)

Meaning	Piłsudski
which	hemanu
what	hemat
what, why, whose	hemata ~ hematu
why (kusu ‘because’)	hemata kusu
wherefore (kusu ‘because’)	hemata ki-kusu
how many	hempak
when, how	hempara
who	inki an-ku(ru)
whither	nakan
what	nex
what (ta ‘LOC, ALL’)	nex-ta
what, where from	nejava
where (ta ‘LOC, ALL’)	nejta
what	nep
what, how	temana
what, how, in what way	temana-ka

an optional locative/allative case marker *ta*. Saru has a different formation based on the interrogative *hemanta* ‘what’ followed by the translative (Shibatani 1990: 36) or mutative (Bugaeva 2012: 476) marker *ne* that derives from the copula. Most forms have a resonance in *h-*. Perhaps, Ainuic thus not only belongs to the group of languages that have what has been called the KIN-interrogative (e.g., Saru *hunna*), but also exhibits K-interrogatives (§6.2.1). However, the presence of the forms *hunna(k)* ‘who’ and *hunnakta* ‘where’ (Batchelor 1905), the latter with locative case marker, suggests that the form underlying both may have been a selective interrogative. Table 5.3 lists some Sakhalin Ainu interrogatives as recorded by Bronisław Piłsudski. For the Sakhalin dialect, Tittel (1922: 77) only mentions a handful of forms that are more or less identical with those listed in Table 5.3. These data clearly show that there are also resonances in *m-* and especially *n-* as well (see also Batchelor 1905).

Vovin (1993) reconstructs four interrogative stems for Proto-Ainuic,  $*gEm=$ ,  $*gu[n]na$ ,  $*in[a]=$ , and  $*nEE=$ , but the situation seems to be much more complicated than that. Altogether he assumes seven interrogatives that are based on these stems as  $*gEm=$  is thought to be the basis for the three different interrogatives  $*gEm=an=ta$  ‘what’,  $*gEm=pa=ra$  ‘which’, and  $*gEm=pak=pE$  ‘how many’ (e.g., Horobetsu *hemanta*, *henpara*, and *henpakpe*), which is in accordance with Cysouw’s (2005) typology and suggests an original meaning ‘which’ or maybe ‘what’. However, there are several problems with Vovin’s reconstructions. Vovin does not comment on the morphology he reconstructs. The use of the equal sign instead of the usual hyphen for morphemes remains unclear as well. Furthermore, it is rather questionable whether an original bilabial nasal *m* should have developed into an *n* followed by a bilabial plosive in all dialects but one. In fact, exactly the opposite development would be expected. Perhaps the same is true for the initial consonant  $*g-$  that in almost all dialects mentioned has the form *h-*. Similarly, except for one dialect, the alleged interrogative  $*nEE=$  actually always has the form *ne*. The stem *ne* is said to mean both ‘who’ and ‘what’, which is rare from a typological perspective, but seems possible (Cysouw 2005; 2007). The interrogative  $*in[a]=$  appears to be mistaken, as there may have been an original initial consonant, e.g. Saru (*h*)*inaan* ‘which’. It may also be noted that Vovin’s (1993) list of cognates is not exhaustive. There is an older but more complete description of interrogatives by Asai (1974: 64f.) that is given in Table 5.4. Figure 5.1 indicates the geographical distribution of the personal interrogatives.

5 Survey of the grammars of questions in Northeast Asia

Table 5.4: Distribution of forms among dialects after Asai (1974: 64f.); 1 = Yakumo, 2 = Oshamambe, 3 = Horobetsu, 4 = Piratori, 5 = Nukibetsu, 6 = Ni-ikappu, 7 = Samani, 8 = Obihiro, 9 = Kushiro, 10 = Bihoro, 11 = Asahikawa, 12 = Nayoro, 13 = Sōya, 14 = Ochiho, 15 = Tarantomari, 16 = Maoka, 17 = Shiraura, 18 = Raichishika, 19 = Nairo, 20 = Kuril, 21 = Chitose

Meaning	Form	Dialects
who	nen	1-3, 7-13
	hunna	4-6, 21
	hunat	20
	naat	14-19
what	nep	1-3, 7-14
	hemanta	3, 4, 6, 21
	hemata	14-19
	hinta	5
	nejanokonejanpi	20
how	nekona	1-3, 9, 10
	nekon	7-13
	mak	4-6
	manak, makanak	21
	temana	14-19
	uiman	20
when	henpara	1-6, 11, 12, 14-16, 18, 21
	henparakanne	17, 19
	henpaki	20
where	nenpara	7-10, 13
	nejta	1-3, 7-13
	nahta	14-19
	hunakta	4, 6, 21
	hunakun	5, 21
	huija	20

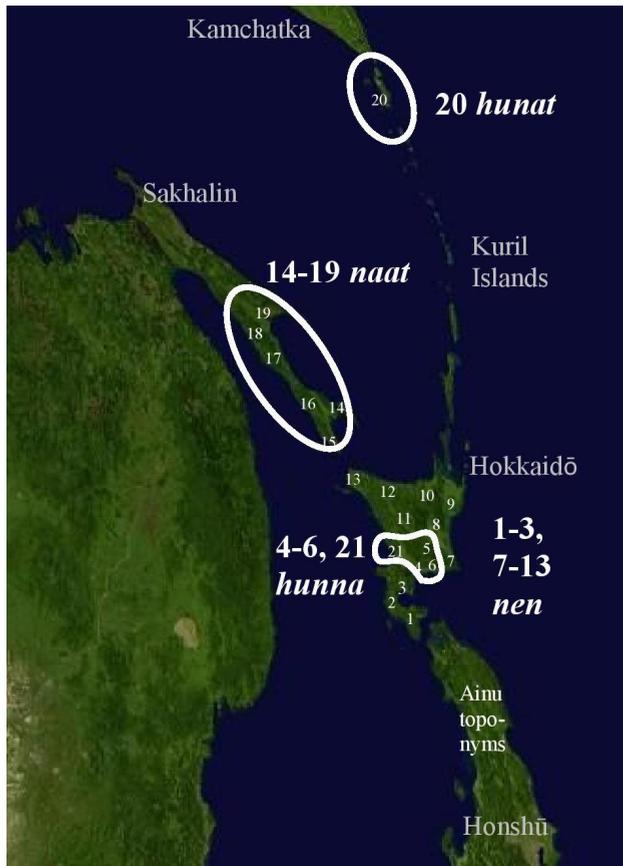


Figure 5.1: Distribution of forms meaning ‘who’ after Asai (1974: 64f.)

## 5.2 Amuric

### 5.2.1 Classification of Amuric

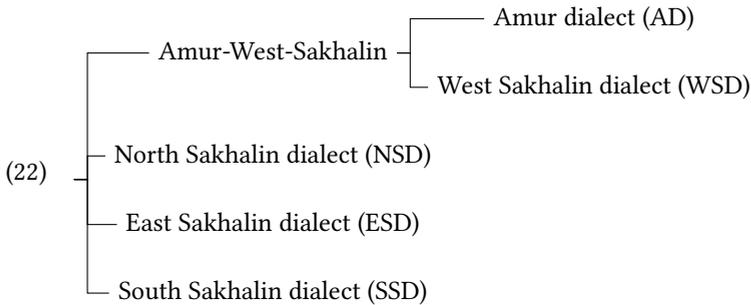
Nivkh is usually considered a linguistic isolate (e.g., Anderson 2006c), but there may be some reason to assume a connection to Chukotko-Kamchatkan languages (Fortescue 2011) (§5.3). Apart from that, there is perhaps enough internal variation to consider it a small language family that will be called Amuric (Janhunen 1996). However, these varieties are traditionally called *dialects* instead of *languages* (Gruzdeva 1998: 7). The relation of these so-called dialects has been characterized by Gruzdeva (1998: 7) as follows:

AD and ESD are rather different: their speakers affirm that they do not understand each other. NSD (or the Shmidt dialect) occupies an in[t]ermediate position between these two. As for SSD (or the Poronaisk dialect), it has essential differences

## 5 Survey of the grammars of questions in Northeast Asia

in phonology, grammar, and vocabulary from the other three dialects, especially from AD.

The Amur dialect has also been spoken on northwestern Sakhalin. Shiraishi (2006) has additionally argued for the existence of a West Sakhalin dialect (WSD) that is different from, but closely related to the Amur dialect (see also Shiraishi & Tangiku 2013). This has not been recognized by Fortescue (2016). In sum, there are the following varieties.



Most examples will be drawn from AD and ESD. The somewhat obscure transcription of some publications has been changed and roughly follows Shiraishi & Tangiku (2013: 203).

### 5.2.2 Question marking in Amuric

According to Gruzdeva (1998: 45), Nivkh makes a distinction between two types of polar question markers. The first type is a suffix that directly attaches to the verb stem and has the form  $-l(o)$  in both Amur and East Sakhalin dialects. The form  $-lo$  is more polite and ceremonious than  $-l$ , which seems to have a more colloquial flavor (Nedjalkov & Otaina 2013: 116).

- (23) Nivkh  
*tʰi ra-l(o)?*  
 2SG drink-FIN.Q  
 ‘Did you drink?’ (Gruzdeva 1998: 45)

- (24) Nivkh (Amur)  
*if pʰri-l(o)?*  
 3SG come-Q  
 ‘Did (s)he come?’ (Nedjalkov & Otaina 2013: 116)

The second type attaches to a finite verbal form or other elements in focus. It has the form  $=l(a) \sim =lo$  in the Amur dialect and the form  $=l(a) \sim =lu$  in the East Sakhalin dialect. It was also written with a hyphen but is reanalyzed as enclitic here. The semantic difference between the two markers, which are perhaps etymologically connected, remains unclear.

(25) Nivkh (Amur)  
*tʃʰi ra-d=la?*  
 2SG drink-FIN=Q  
 ‘Did you drink?’

(26) Nivkh (East Sakhalin)  
*tʃʰi ra-d=lu?*  
 2SG drink-FIN=Q  
 ‘Did you drink?’ (Gruzdeva 1998: 45)

Polar and focus questions have the same marker that attaches to the verb in the former and to the element under focus in the latter.

(27) a. Nivkh (Amur)  
*itik pʰri-dʒ=la?*  
 father come-IND=Q  
 ‘Has father come?’  
 b. *itik=la pʰri-dʒ?*  
 father=Q come-IND  
 ‘Is it *father* who has come?’ (Nedjalkov & Otaina 2013: 124)

Content questions may be unmarked if they have a special intonation that was left unspecified by Gruzdeva (1998: 46). Otherwise they have a question marker different from that for polar questions (Amur =*ŋa*, =*at(a)*, East Sakhalin =*ŋa*, =*ŋu*, =*ara*). The markers may either attach to the verb or the interrogative (phrase). They have been reanalyzed as enclitic here. Interrogatives remain *in situ*.

(28) Nivkh (East Sakhalin)  
 a. *tʃʰin tʰamdʒi pʰ-vo-ux tʰamdʒi ŋa turʰ-pirʰk*  
 2PL what.kind REFL-village-LOC what.kind animal meat-?only  
*ʃi-tʰa-d-ʒun?*  
 eat-HAB-IND-PL  
 ‘What kinds of animal meat do you eat in your village?’ (Chae 2013: 132)  
 b. *tʃʰi tʰa.k-toχ vi-d=ŋa?*  
 2SG where-DAT go-IND=Q  
 ‘Where are you going (roughly)?’ (Gruzdeva 2008: 182)

(29) Nivkh (Amur)  
 a. *aŋ pʰri-dʒ=at?*  
 who come-IND=Q  
 ‘Who came?’

5 Survey of the grammars of questions in Northeast Asia

- b. *tʃ<sup>ni</sup> sidʒ=ŋa j-isru-dʒ?*  
 2SG what=Q OBJ-pursue-IND  
 ‘Whom do you pursue?’ (Gruzdeva 1998: 46)

- (30) Nivkh (Amur)  
*sidʒ jivɣ=ŋa jiv-dʒ?*  
 what person=Q have-IND  
 ‘What (kind of) man is (here)?’ (Nedjalkov & Otaina 2013: 124)

The existence of separate and overtly marked polar and content question markers seems to have been adopted by the Tungusic language Uilta (§5.10.2).

No clear examples for tag questions and only one example for a negative alternative question have been found. The analysis of this example from von Glehn (Grube 1892: 31) remains partly obscure for me but is sufficiently clear to show that there is no disjunction and that each alternative takes a marker *lo*. In the Amur dialect this may either correspond to the enclitic =*l(a)* ~ =*lo* or to the suffix -*l(o)*. However, Nedjalkov & Otaina (2013: 125, 209) mention a suffix -*lu* found in the Amur dialect, misleadingly called “particle” despite being given with a hyphen, that seems to have dubitative meaning and marks indirect alternative questions. Given that it may also have the form -*lo*, it seems possible that this is the form recorded by von Glehn.

- (31) Nivkh (Amur)  
*[tu-ni-dʒ-lu                      qa-ni-dʒ-lu]                      p<sup>h</sup>anp<sup>h</sup>ara-r*  
 go.upstream-FUT-N-DUB go.downstream-FUT-N-DUB not.know-CVB.NAR.3SG  
*hum-dʒ.*  
 be-IND  
 ‘He does not know [whether to go upstream or downstream].’ (Nedjalkov & Otaina 2013: 209)

An etymological connection to the other two question markers seems likely but to my knowledge there has not been an investigation of this topic. The same marker also appears in indirect polar questions (Nedjalkov & Otaina 2013: 220) and content questions (see 33a,b below). This quite clearly shows that it should be kept apart from the actual question markers. On the contrary, it may be a marker for indirect questions, exclusively. Rhetorical questions in Nivkh are marked with -*rla* ~ -*tla*.

- (32) Nivkh (Amur)  
*if p<sup>h</sup>ri-rla?*  
 3SG come-Q  
 ‘Did (s)he really come?’ (Nedjalkov & Otaina 2013: 116)

A special marker that is said to expect a positive answer and thus perhaps comes close to a question tag is (probably sentence-final) <*y*> as recorded by von Schrenck (Grube 1892). Austerlitz (1956: 262) mentions a marker =*ii*, reanalyzed as enclitic here, that he

translates as ‘isn’t it?’ and it might be the same as <y>, e.g. *ηav=ii?* ‘a sparrow’s nest, isn’t it?’. The Tungusic language Uilta (§5.10.2) not only has a content question marker =*ga* ~ =*ka* that most likely derives from Nivkh =*ηa* (§3.1), but also has a polar question marker =(y)*i* that could to stem from this enclitic in Nivkh.

Table 5.5: Summary of question marking in Amuric.

	PQ	CQ	AQ	FQ
AD	V=l(a)/=lo	= <i>ηa</i> , =at(a)	2x =l(a)/=lo, ? 2x =lu/=lo	FOC=l(a)/=lo
ESD	V=l(a)/=lu	= <i>ηa</i> , = <i>ηu</i> , =ara	?	FOC=l(a)/=lu

Slightly adjusting Fortescue’s (2016: 79, 172) reconstructions, Proto-Amuric must have had the question markers \*=*la* ~ =*lo*, \*=*rla* ~ =*rlo*, \*=*ηa*, =*ata*, and \*=*i* with somewhat unclear distribution.

### 5.2.3 Interrogatives in Amuric

Descriptions of interrogatives in Nivkh are usually insufficient, especially for the South and North Sakhalin dialects. Table 5.6 shows those forms collected by Mattissen (2003) and Fortescue (2016) to which WSD data has been added (Shiraishi & Tangiku 2013). The Amur and West Sakhalin dialects have a resonance in *ř*~ and the East Sakhalin dialect in *t*<sup>h</sup>~ that go back to the same origin. Interrogatives meaning ‘what’ and ‘when’, and, except for ESD, also the interrogative meaning ‘who’ do not participate in this resonance. The resonance has been recorded as *š*~ by von Schrenck and as *s*~ by von Glehn (Grube 1892). For example, von Schrenck had a form *ša*- ‘which, what kind of’ (AD *řa*-) as well as its regular locative form *ša-in* ‘where’ (AD *řa-in*, Fortescue 2011: 144).

Fortescue (2016: 111) speculates that AD *aη* derives from *nar-ηa* ‘who-Q’. If correct, a typological parallel can be found in Korean (§5.8.3). ESD *t<sup>h</sup>au-nt/-d* ‘who’ is perhaps a secondary innovation based on the selective interrogative. Interestingly, almost all listed interrogatives are monosyllabic. But there are some longer forms as well, as the following two examples from the Amur dialect illustrate.

(33) Nivkh (Amur)

- a. [*jagut imη p<sup>h</sup>ri-dž-lu*] if *p<sup>h</sup>anp<sup>h</sup>ara-dž*.  
how.3PL 3PL come-IND-DUB 3SG not.know-IND  
‘He does not understand [how they came there].’
- b. [*jagur p<sup>h</sup>ri-dž-lu*] *p<sup>h</sup>anp<sup>h</sup>ara-dž*.  
how.3SG come-IND-DUB not.know-IND  
‘(He) does not understand [how (he) came (there)].’ (Nedjalkov & Otaina 2013: 220)

5 Survey of the grammars of questions in Northeast Asia

Table 5.6: Nivkh interrogatives according to Mattissen (2003: 14) and Fortescue (2016: passim), WSD according to Shiraiishi & Tangiku (2013: 206); not all variants listed

Proto-Amuric	AD	WSD	ESD	NSD	SSD
*nar 'who'	aŋ ~ aŋ	aŋ	nar ~ nař, t <sup>h</sup> au-nt/-d	nar ~ nař	nat
*tu-nt 'what'	si-dʒ	si-tʃ ~ si-dʃ	ru-(n)t/-d	ru-t, řu-t, su-t	ru-nt, lu-nt
*ta-nt 'which'	řa-dʒ ~ t <sup>h</sup> a-dʒ	? <sup>1</sup>	t <sup>h</sup> a-d	?	?
*taŋʒ ~ *taŋr 'how much/ many'	řa-ŋs	řa-ŋs	t <sup>h</sup> a-ŋs ~ t <sup>h</sup> a-gs	řa-ŋspaklu 'some'	t <sup>h</sup> a-ŋk ~ řa-nkř 'some'
*ta- 'where'	řa-r	řa-ŋ ~ řa-g	t <sup>h</sup> a-s	?	řa-k, t <sup>h</sup> a-k
*ayr 'when'	iyr	iyr	ayř ~ iyř	<b>irja</b>	axř
*ja-(ni-ŋ) 'how, why'	ja-ŋu-t/-r	jaŋ-gu-ni-tʃ	ja-ŋř ~ ja-nř	ja-na-gu-t	ja-ni-ŋ, ja-n-ř, ja-ni-g

Also observe the dubitative suffix *-lu* used for indirect questions presented in §5.2.2. In *jagur ~ jagut* the element *-r* (2SG, 3SG) ~ *-t* (1SG, 1PL, 2PL, 3PL) is the narrative converb marker that is also part of the rhetorical question marker *-r-la ~ -t-la* previously noted (Nedjalkov & Otaina 2013: 40). The forms also contain an old causative marker *-ku ~ -yu ~ -gu ~ -xu* that apparently has mostly lost its function (Nedjalkov & Otaina 2013: 42). Apparently, Gruzdeva (1998) and Mattissen (2003) do not mention any of these forms, but they have been listed as *ja-ge-r* (von Schrenck), *jaŋ(-o-r)* (von Glehn), *ja-g-r* (Seeland), and *jan-g-r* (Lebedew) by Grube (1892). As in two of these examples, AD and WSD sometimes contain a consonant *-ŋ* which—Nedjalkov & Otaina (2013: 87) speculate—might be a dialectal difference. According to Fortescue (2016: 81), the *-ŋ* could be a participle form. Table 5.7 shows the paradigm of these forms as can be reconstructed with the help of different descriptions.

But according to Shiraiishi & Tangiku (2013: 206) there are also some longer forms such as WSD *jaŋ-gu-ni-tʃ* 'how'. The WSD suffix *-tʃ* is the same as AD *-dʒ* 'IND' that attaches to what appears to be the future marker *-ni* (Nedjalkov & Otaina 2013: 209) or perhaps the verb *-ni ~ -nu* 'to do' as in SSD *ja-ni-ŋ* (Fortescue 2016: 81). Nedjalkov & Otaina (2013: 369) mention in addition an AD form *jaar* 'why' that must be related to these forms

<sup>1</sup>Given the parallel in the AD and ESD, one may assume that the WSD has the form *řa-tʃ ~ řa-dʃ* 'which'.

Table 5.7: Simple AD and WSD interrogative paradigms of the form meaning ‘how’ (Nedjalkov & Otaina 2013: 40, 220, Shiraishi 2006: 65, Shiraishi & Tangiku 2013: 206)

	SG	PL
1	ja(ŋ)-gu-t	ja(ŋ)-gu-t
2	ja(ŋ)-gu-r	ja(ŋ)-gu-t
3	ja(ŋ)-gu-r	ja(ŋ)-gu-t

but has a long vowel and lacks the causative suffix (see Fortescue 2016: 81 for additional variants). According to Mattissen (2003: 238) the stem *ja-* (optionally with a derivation *ja-γa-* not encountered thus far) actually means ‘to do what’. The forms *ja-* as well as *ja-ɸo* may also be employed as an attribute, e.g. AD *ja-nivx* ‘what person’, *ja-ɸo-dəf* ‘what kind of house’. These patterns are extremely similar to Mongolic (\**ya-xu/n* ‘what’, \**ya-xa-* ‘to do what’, §5.8.3) and Tungusic (\**ja-(kun)* ‘what’, \**ja-* ‘to do what’, §5.10.3).<sup>2</sup> Possibly, the Nivkh forms are Tungusic loans that in turn derive from Mongolic. The converbal origin of forms meaning ‘how’ or ‘why’ might also suggest a connection with Mongolic or Tungusic. Within Nivkh there are completely parallel forms in the demonstratives, e.g. AD *ho-(ɸo)-* ‘be like that, do thus’, *ho(ŋ)-gu-r/t* ‘thus, in that way’ etc. (Nedjalkov & Otaina 2013: 87f.).

Suffixes in the locative (AD *řa-r*, ESD *t<sup>h</sup>a-s* ‘where’) and the quantitative interrogatives (AD *řa-ŋs*, ESD *t<sup>h</sup>a-ŋs* ~ *t<sup>h</sup>a-gs* ‘how much/many’) have parallels in spatial expressions and demonstratives, cf. AD *tu-r* ‘here’, *hu-r* ‘there’, *tu-ŋs* ‘this much’, *hu-ŋs* ‘that much’ (Gruzdeva 1998: 26f., 36), ESD *tu-s*, *hu-s*, and *tu-nks*, *hu-nks* with a slightly different form (Gruzdeva 2008: 170). Mattissen (2003: 14) furthermore mentions AD *řa-kr* ~ *t<sup>h</sup>a-kr* ‘where’ that has a suffix also known from spatial expressions and demonstratives, e.g. ESD *tu-kř* ‘here’, *hu-kř* ‘there’ (Gruzdeva 2008: 181). The difference between *-s* and *-kř* is that the former designates a precise and the latter a non-precise location (Gruzdeva 2008: 178). Another suffix *-nx* roughly patterns with the latter in meaning, e.g. ESD *t<sup>h</sup>a-nx* ‘where’ (Gruzdeva 2008: 184). It is possible to attach a case marker such as the dative to the locative forms, e.g. ESD *t<sup>h</sup>a-s-toχ*, *t<sup>h</sup>a-k-toχ* ‘where to’ (Gruzdeva 2008: 179, 182). Thus, similar to Tungusic the forms meaning ‘where’ are derived from the selective interrogative (AD *řa-dʒ* ~ *t<sup>h</sup>a-dʒ*, ESD *t<sup>h</sup>a-d*).

The forms meaning ‘what’ may be analyzed as a stem and the nominalizer (indicative) \**-nt* > AD *-dʒ*, ESD *-nt* ~ *-(n)d* etc. (Fortescue 2011: 1366). The same element is present in the selective interrogative and ESD *t<sup>h</sup>au-nt*, *t<sup>h</sup>au-d* ‘who’, as well as some demonstratives (Table 5.8). Notice that von Schrenck recorded the Amur dialect form meaning ‘what’ as *si-č* ~ *si-nč* (Grube 1892), which preserves a nasal that is also present in ESD *ru-d* ~ *ru-nt* ‘what’.

<sup>2</sup>According to Nedjalkov & Otaina (2013: 209) and Fortescue (2016: 81), the initial *j-* is a third person singular marker—a hypothesis first proposed by Jakobson—while *a-* is the actual interrogative verb meaning ‘to do what’. But the connection with Tungusic and Mongolic makes this very unlikely.

## 5 Survey of the grammars of questions in Northeast Asia

Table 5.8: Amuric demonstratives and interrogatives “indicating a person or an object” (Gruzdeva 1998: 26ff.)

	AD	ESD
DEM.PROX	ti-dʒ	tu-d ~ tu-nt ~ ti-nt
DEM.DIST 1	hi-dʒ	hu-d ~ hu-nt ~ hi-nt
DEM.DIST 2	a-dʒ	ahu-d ~ eh-u-d ~ ehi-nt
DEM.DIST 3	aehi-dʒ	aix-nt
DEM.invisible	ku-dʒ	ku-d ~ ku-nt
what	si-dʒ	ru-d ~ ru-nt
which	řa-dʒ ~ t <sup>h</sup> a-dʒ	t <sup>h</sup> a-d
who	-	t <sup>h</sup> au-d ~ t <sup>h</sup> au-nt

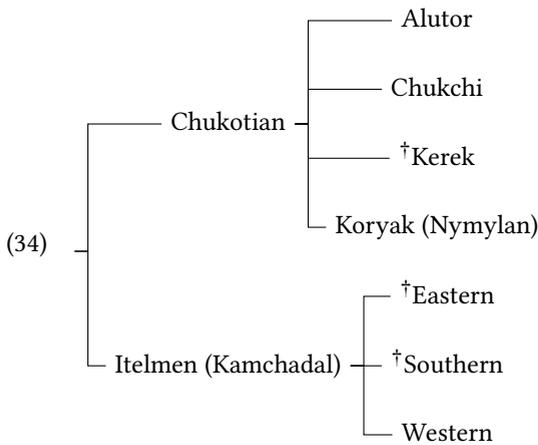
Demonstratives with the suffix may take number and case markers (e.g., AD *ti-dʒ-Ø-yir* ‘this-IND(-SG)-INST’), without, they may function as attributive forms (e.g., AD *ti urk* ‘this night’). Perhaps a similar situation can be observed for the interrogatives *t<sup>h</sup>amdʒi* ‘what kind of’ (Chae 2013: 135) versus *t<sup>h</sup>amdʒi-d* ‘how’ (Fortescue 2011: 1372) in the ESD (similar to *ja-dʒ ~ ja-* in AD, Mattissen 2003: 238).

Fortescue (2011: 1371) assumes that Nivkh *t<sup>h</sup>a-/řa-* is related to \**ǰæq* in Proto-Chukotko-Kamchatkan (e.g., Chukchi *räq*, Alutor *taq*). He reconstructs a common proto-form for both as \**tA(q)-* (§5.3.3). But as long as the hypothetical language family is not accepted by a majority of scholars, this must be treated with caution. Two interrogatives from Nivkh may have found their way into the Tungusic language Uilta (§5.10.3). The Uilta materials collected by Bronisław Piłsudski contain the two forms *nuulú* ‘whither’ and *sádo* ‘where’ (Majewicz 2011: 388, 430). The second interrogative also has the form *saa* ‘where’ with a long vowel and is most likely a loan from West Sakhalin Nivkh *řa-g* ‘where’ (cf. Ikegami 1997; Pevnov 2009: 122). Note that von Glehn recorded several forms starting with *s-* (Grube 1892). Allegedly, the ESD also has the forms *nu-nt ~ nu-d* ‘what’. Fortescue (2011: 1372) speculates that these forms are actually indefinites and may contain a contracted form of the noun *nə-* ‘thing’. But if Uilta *nuulu* is indeed from Nivkh, it must be connected somehow to this form in the East Sakhalin dialect.

## 5.3 Chukotko-Kamchatkan

### 5.3.1 Classification of Chukotko-Kamchatkan

Chukotko-Kamchatkan (or Luoravetlan) is a small family that includes five languages in two different branches (Fortescue 2003: 51f.; Anderson 2006a).



Itelmen formerly consisted of three different languages or dialect groups, of which all but the Western group have already become extinct. Kerek disappeared during the 1990s. Recently, it has been proposed that Amuric (§5.2) may be distantly related to Chukotko-Kamchatkan (Fortescue 2011), but this hypothesis remains unproven.

### 5.3.2 Question marking in Chukotko-Kamchatkan

Given the lack of data on other question types, the following will focus primarily on polar and content questions. *Alutor* marks polar questions by means of probably rising intonation and an optional question particle. Unlike most other languages treated in this study, the particle does not stand sentence-finally but initially, which, except for some Indo-European languages (§5.9.2), represents a stark contrast with NEA (Chapter 6). Content questions have no question particle.

(35) *Alutor*

- a. *matka ta-lyu-ŋi?*  
 Q POT-kill.wild.reindeer-POT.PFV[3SG.S]  
 ‘Will he kill a wild reindeer?’
- b. *miyya iv-i?*  
 who.ABS.SG say-3SG.S[PFV]  
 ‘Who said (that)?’ (Nagayama 2011: 293, 294)

In all Chukotko-Kamchatkan languages, interrogatives seem to take sentence initial position, which likewise differentiates them from the rest of NEA. Interestingly, the initial question particle itself looks similar to Chukotko-Kamchatkan interrogatives starting with *m-* (see §5.3.3). Fortescue (2005: 416) translates *matka* as ‘or’ and lists it with forms such as Chukchi *mec-* ‘somewhat’. While the exact derivation remains unexplained, there is also a Koryak form *met(‘)ke* ‘or’ that appears to be a direct cognate of *Alutor matka*. Content questions are likewise unmarked in Koryak.

5 Survey of the grammars of questions in Northeast Asia

- (36) a. Koryak  
*met'ke jenny e-jem-ke?*  
 Q maybe NEG-come-CIRC  
 'Perhaps (she) does not come?'  
 b. *meki ib-i?*  
 who.ABS.SG say-3SG.S[PFV]  
 'Who said (that)?' (Zhukova 1997: 51)

It seems that in Chukchi and Kerek both polar and content questions are generally unmarked.

- (37) Kerek  
 a. *mil'ej jallaju-ŋi-n?*  
 gun take-PRF-3SG.O  
 'Will you take this gun?'  
 b. *maki jatč-i?*  
 who.ABS.SG come-PRF  
 'Who came?' (Volodin 2001: 156, 157)

- (38) Chukchi  
 a. *koolo enmec ye-ŋjew-iyət?*  
 INTJ already PF-awaken-2SG  
 'My goodness, you're up already?'  
 b. *tite ŋan ŋotqen n-ə-qit-ə-qin?*  
 when DEICT DEM.3SG.ABS HAB-E-freeze-E-3SG  
 'When does it freeze there?' (Dunn 1999: 86, 72)

Chukchi furthermore has an element *ətlon*, glossed as a question marker, that appears in both polar and content questions and was translated as 'on earth', i.e. it adds a certain emphasis. It may also fuse with the interrogative *əmi* 'where' to form the more complex emphatic interrogative *əmitlon* 'where on earth' (Dunn 1999: 289f.). Its syntactic position is not absolutely clear, however, but seems to be relatively free.

- (39) Chukchi  
*anə kəkəl, ətlon i'am, req-ə-l'et-ə-rko:n?*  
 so INTJ Q why what-E-DUR-E-PROG.VOC  
 'Oh my! Why, what on earth are you doing?' (Dunn 1999: 55)

It does not seem to be a true question marker, but nevertheless appears in interrogative contexts. Functional equivalents can be found in Yiddish (§5.5.2.2) and Tundra Nenets (§5.12.2).

Polar questions in Itelmen have final rising intonation but otherwise are identical to equally unmarked content questions.

- (40) Itelmen  
*kni-n gitkinej çi-ze-n?*  
 PP.2SG-POSS brother be.available-PRS-3SG  
 ‘Do you have a brother?’ (Georg & Volodin 1999: 214)

Interrogatives in content questions optionally take a suffix *-s*, which is said to be a question marker that expresses additional emphasis.

- (41) Itelmen
- a. *k'e ç'e'ł-en k'oł-ki?*  
 who COM.V-2SG.3SG come-INF.V  
 ‘With whom have you come?’
- b. *manke-s kamma-n mnił k-meç-knen?*  
 where-Q PP.1SG-POSS all INF.III-disappear-CIRC  
 ‘Where have all my (people?) gone?’ (Georg & Volodin 1999: 134, 214)

Itelmen is the only Chukotko-Kamchatkan language for which descriptions of focus questions are available to me. They follow an intriguing pattern that has a variable personal marker on the verb.

- (42) Itelmen
- a. *isx-enk n-zal-at-in kza kama-nk?*  
 father-LOC IMPRS-give-FUT-2SG.OBJ.O you me-DAT  
 ‘Will father give *you* to me?’
- b. *isx-enk n-zal-at-um kza kama-nk?*  
 father-LOC IMPRS-give-FUT-1SG.OBJ.IO you me-DAT  
 ‘Will father give you *to me*?’ (Bobaljik & Wurmbrand 2002: 3)

In this example, either the direct or the indirect “object” are represented with an agreement marker on the verb. The presence of the marker expresses the focusing of the respective constituent.

Chukotko-Kamchatkan languages have a strong interaction of **imperatives** and question marking, which is yet another untypical feature for NEA. For example, Nedjalkov (1994: 325) mentions the interesting fact that imperative verb forms in Chukchi may appear in content questions where their meaning changes to marking future tense.

- (43) a. Chukchi  
*myn-le-rkyn?*  
 IMP.1PL-go-IPFV  
 ‘Let us fly!’
- b. *minky.ty myn-le-rkyn?*  
 over.where IMP.1PL-go-IPFV  
 ‘Over what place shall we fly?’ (Nedjalkov 1994: 325)

Georg & Volodin (1999: 171) claim that imperatives in Itelmen may also have a future and prospective meaning, but this does not appear to be restricted to questions. The phenomenon in Chukchi has a more straightforward parallel in the more closely related language Kerek, for which Volodin (2001: 158) noted the following phenomenon:

Interrogative sentences in Kerek are often viewed as a special type of imperative utterances that presuppose a speech response. Any interrogative sentence can be interpreted as a reduced imperative sentence of the type “Tell (answer) me, if...”. This view may be confirmed by the strong formal ties existing between imperative and interrogative meanings demonstrated by Chukchi-Koryak (and Chukchi-Kamchatkan) languages.

In both Kerek and Chukchi the imperative markers in questions exhibit an additional modal overtone such as ‘can’ or ‘must’ (Volodin 2001: 157).

- (44) Kerek  
*manka nə-xaxau-n?*  
 why IMP.3SG-go-3SG.S  
 ‘Why does he have to go?’ (Volodin 2001: 156)

The imperative marker is not obligatory, however, and as in Chukchi all examples provided by Volodin are content questions. Whether this feature is shared by Alutor and Koryak remains unclear for now. Interestingly, interrogative morphology in the adjacent Yukaghiric languages (see §5.14.2) as well as in Central Alaskan Yupik (§5.4.2) is also restricted to content questions. See also §5.10.2 on Even, a Tungusic language that had contact with Chukotko-Kamchatkan and exhibits the use of imperative forms in questions as well.

The marking of questions in Chukotko-Kamchatkan summarized in Table 5.9 exhibits no similarities to Amuric or to most of NEA, for that matter.

Table 5.9: Summary of question marking in Chukotko-Kamchatkan

Language	PQ	CQ
Chukchi	-	-(IMP-V)
Alutor	#matka	-
Kerek	-	-(IMP-V)
Koryak	#met’ke	-
Itelmen	-	(#INT-S)

### 5.3.3 Interrogatives in Chukotko-Kamchatkan

Several Proto-Chukotko-Kamchatkan (PCK) interrogatives have been reconstructed by Fortescue (2005). Table 5.10 lists them with cognates from all five languages, but not all

variants and only singular forms are shown. Each language has some additional forms, e.g. *laʹlsxeʹn* ‘how much/many’, *manke* ‘whence, how’, *manxʹal* ‘whither’, *əŋqa* ‘what’, and *əŋqan-kit* ‘what-CAUS > why’ in Itelmen (Georg & Volodin 1999: 136, passim), *maŋ-ki*, *maja* ‘where’, *maŋ-kət(iŋ)* ‘whence’, *maŋ-kepəŋ* ‘whence, along where’, *maŋ-inʹas* ‘how many, how long’, and *taʹɬər* ‘how much’ in Alutor (Nagayama 2011: 293f.), and *ʹemi* ‘where’, *iʹam* ‘why’, *mik-ə-ne* ‘whither’, *tʹer* ‘how much/many’ etc. in Chukchi (Dunn 1999: 66, passim). The most important resonance of Chukotko-Kamchatkan languages is *m~*.

Table 5.10: Proto-Chukotko-Kamchatkan (PCK) interrogatives and their cognates in individual languages according to Fortescue (2005: 56, 173, 175ff., 287) and Dunn (1999; 2000)

Meaning	PCK	Chukchi	Kerek	Koryak	Alutor	Itelmen
what kind	*mæŋjin	məŋjin	maŋjin	məŋjin	maŋjin	min
who (ABS.SG)	*mikæ	mik(ə)-	maki	meki, maki (Kamen)	miŋŋa	kʹe
where	*miŋ(kə)	miŋkə	miŋkiil (ALL)	miŋkə	miŋkə	maʹ, mank
how	*miŋkəði	miŋkəri	miŋkii	miŋkəje	?maŋkət	?mank
what	*ðæq-	req-/ceq-	jaq-	jeq-	taq-, teq- (Palana)	saq
when	*titæ	tite	sita	tite	tita	itʹe

Fortescue (2005: 263, 282) reconstructs, furthermore, Proto-Chukotian (PC) stems that lack a cognate in Itelmen, i.e. PC *\*Ræmi* ‘where’ (Chukchi *ʹemi*, Kerek *Xam*, and Koryak *hemmi*, Alutor -) and PC *\*tæRər* ‘how much’ (Chukchi *tʹer*, Kerek *tʹaj*, Koryak *teri*, and Alutor *tarər*). Itelmen likewise exhibits interrogatives without clear equivalents in Chukotian such as one meaning ‘what’ (Eastern *nkc*, Southern *nakxej*, and Western *əŋqa*, Fortescue 2005: 399). Fortescue (2011: 1372) compares PCK *\*ðæq-* ‘what’ with Nivkh *tʰa-/řa-* (§5.2.3) and tentatively reconstructs PCKA *\*tʰ(q)-*. However, this reconstruction is still too speculative, given that the genetic connection between the two families has not been proven beyond doubt. This stem in Chukotko-Kamchatkan cannot only have nominal but also verbal properties.

- (45) Alutor  
*yəttə      taq-ətkən?*  
 2SG.ABS.SG what-IPFV[2SG.S]  
 ‘What are you doing?’ (Nagayama 2011: 294)

5 Survey of the grammars of questions in Northeast Asia

- (46) Koryak (Kamenskoye)  
*ni-ya'q-iyi?*  
 HAB-what-2SG  
 'What are you doing?' (Bogoras 1922: 730)
- (47) Koryak  
*n-re'q-iyit?*  
 HAB-what-2SG  
 'What are you doing?' (Bogoras 1922: 730)
- (48) Chukchi  
*nə-req-iyət?*  
 HAB-what-2SG  
 'What are you doing?' (Dunn 1999: 368)

Chukchi earlier made a characteristic difference between *req-* as used by men and *ceq-* as used by women (Kämpfe & Volodin 1995: 8). But this is just the effect of a more general pattern in which women pronounced *r* as *c* that seems to have been lost by now (Dunn 2000). Another language in Northeast Asia that makes some distinctions between the grammar of questions of women and men is Japanese (§5.6.2). Similar to Ket (§5.13.3), interrogatives can be incorporated into the verb. When incorporated the meaning of *req-/raq-* ~ *r'e-/r'a-* changes from 'what' to 'why'.

- (49) Chukchi
- a. *raq-etə nə-wetgawe-g'ət?*  
 what-DAT PRS-speak-2SG.S  
 'Why do you speak?'
  - b. *nə-raq-ə=wetgawe-g'ət?*  
 PRS-what-E=speak-2SG.S  
 'Why do you speak?'
  - c. *r'a-etə ηətə-g'ət?*  
 what-DAT come-2SG.S  
 'Why did you come?'
  - d. *r'a=ηətə-g'ət?*  
 what=come-2SG.S  
 'Why did you come?' (Spencer 1995: 457, from Skorik)

As examples (49a) and (49c) illustrate, the meaning 'why' is otherwise expressed with the dative form of the interrogative. See §5.8.3 and §5.10.3 for a somewhat similar development in Khorchin and Manchu.

Interrogatives in Chukotko-Kamchatkan languages have elaborated paradigms (see Nagayama 2011: 293f. on Alutor; Bogoras 1922: 726ff. on Koryak; Georg & Volodin 1999:

134-136 on Itelmen). In Chukchi the paradigms correspond to the second [+HUM] and first declension [+/-HUM] of nouns, respectively (Table 5.11). In order to make clear the distinction found in the second declension into collective suffixes on the one hand and number/case suffixes on the other, the sign  $\emptyset$  indicates which of the markers is absent. The layering of suffixes follows the order V-COLL-NUM/CASE. The first declension has no collective suffixes. Locative interrogatives and demonstratives have parallel paradigms (Dunn 1999: 286f.), e.g. *ɲut-ku* ‘DEM.PROX-LOC’, *ɲen-ku* ‘DEM.DIST-LOC’, and *miɲ-ke* ‘where-LOC’. The ablative (*meɲ-qo(rə)*) and allative (*miɲ-kəri*) have the same forms throughout.

Table 5.11: Chukchi interrogative paradigms according to Kämpfe & Volodin (1995: 87)

Glossing	who (2nd decl.)	what (1st decl.)
stem	mik-	req-
ABS.SG	meɲi- $\emptyset$ -n	r <sup>ɲ</sup> etnyt- $\emptyset$
ABS.PL	miky- $\emptyset$ -nti	r <sup>ɲ</sup> etnyt-et
LOC-ERG (-COLL)	miky-ne- $\emptyset$	req-e (INST-ERG)
LOC-ERG (+COLL)	miky-ryk- $\emptyset$	req-yk (LOC)
ABL (-COLL)	mek- $\emptyset$ -gypy, (meky-na-jpy)	r <sup>ɲ</sup> a-/raq-gypy
ABL (+COLL)	meky-r-gypy	-
ALL (-COLL)	meky-na- $\emptyset$ (/ -gty)	raq-ety
ALL (+COLL)	meky-ryk-y	-
ORIENT (-COLL)	miky- $\emptyset$ -gjit	reqy-gjit
ORIENT (+COLL)	miky-ry-gjit	-
DESIG (+/-COLL)	miky- $\emptyset$ -ny	req-y

In Alutor, participle forms of the interrogative verb may take case markers as well.

- (50) Alutor  
*ənɲin taq-ə-lʔ-u qa paninalʔ-u?*  
 well what-E-PTCP-ABS.PL EMPH ancestor-ABS.PL  
 ‘Well, what did (our) ancestors do?’ (Nagayama 2016: 133)

Predicatively used interrogatives can also take person and number markers.

- (51) Alutor  
*mik-ine-yət ɣəttə un<sup>i</sup>un<sup>i</sup>u-jyət?*  
 who-POSS-2SG.PRED 2SG.ABS child-2SG.PRED  
 ‘Whose child are you?’ (Nagayama 2016: 121)

Unlike Chukchi or Itelmen, but similar to Aleut (§5.4.3), Alutor and Koryak not only have plural but also dual forms.

In sum, Chukotko-Kamchatkan interrogatives deviate strongly from other NEA languages. No K-interrogatives are present and only Itelmen *k'e* has been tentatively classified as a KIN-interrogative, although it likely derives from what has been reconstructed as PCK \**mikæ*. Complex paradigms with sandhi effects, ergative marking, dual number (e.g., Koryak *ma'ki* 'ABS.SG', *ma'kinti* 'ABS.DU', *maku'wyi* 'ABS.PL', Bogoras 1922), and incorporation set Chukotko-Kamchatkan apart from most other languages in NEA. However, ambivalent interrogative stems meaning '(to do) what' are shared with Tungusic, Eskaleut, and Samoyedic. Especially Itelmen exhibits an opaque interrogative system that resists any synchronic attempt for analysis. An exhaustive diachronic analysis can only be accomplished by experts on the language.

## 5.4 Eskaleut

### 5.4.1 Classification of Eskaleut

The Eskimo-Aleut or Eskaleut language family may be classified as in Figure 5.2 (Berge 2006; 2010; Fortescue 2013; and especially Fortescue et al. 2010: xiif.).

Languages spoken in Northeast Asia are signaled with an asterisk, but for the purpose of better understanding, Central Alaskan Yupik will be included in the discussion as well. For a more fine-grained classification of subdialects see Fortescue et al. (2010: xiif.). The primary split is between Aleut on the one hand and Eskimo on the other. Eskimo itself falls into two main branches, Yupik and Inuit. However, Sirenik(ski)—usually considered a part of Yupik—could possibly form a third branch of Eskimo (Fortescue et al. 2010: x). In general, the Aleut branch must be considered the most aberrant member of the family. Aleut historically formed a dialect continuum with linguistic diffusion from east to west but only three main dialect groups are sufficiently attested (Bergsland 1997: 14). Copper Island or Mednyj Aleut is a truly mixed language that contains a large number of Russian elements, including verbal morphology (e.g., Comrie 1981: 253; Golovko & Vakhtin 1990; Sekerina 1994; Golovko 2003; Vakhtin 1998), but is classified with other Aleut dialects here.

### 5.4.2 Question marking in Eskaleut

Eskaleut languages are famous for their interrogative mood, perhaps because of the well-known description of questions in West Greenlandic by Sadock (1984), but this is not present in all Eskaleut languages. **Aleut** has a mobile question particle, *hi(i)'* ~ *ii'* with final glottal stop in the Eastern dialect and *ii* in Attuan and Atkan (Bergsland 1997: 82) that marks polar and focus questions. It is reanalyzed here as enclitic because it freely attaches to the element in focus. As expected, the finite verb is focal in polar questions.

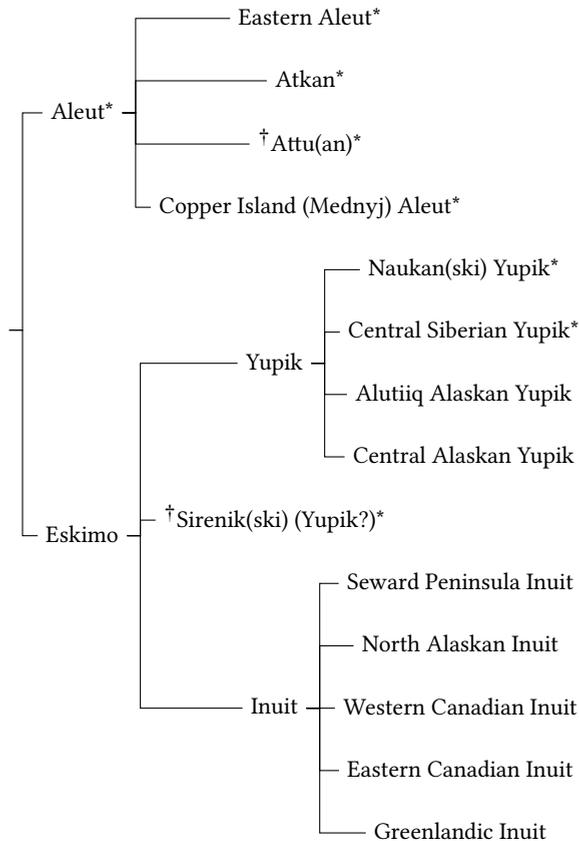


Figure 5.2: Classification of Eskaleut.

## (52) Aleut (Atkan)

- a. *qilagan piitra-x̂ hla-x̂ tuga-l saĝa-na-x̂=ii?*  
 yesterday PN-ABS.SG boy-ABS.SG hit-?CVB AUX-REM-SG=Q  
 ‘Did Peter hit the boy yesterday?’
- b. *qilagan piitrax̂ hla-x̂ tugal=ii saĝana-x̂?*  
 ‘Did Peter really *hit* the boy yesterday?’ (Bergsland 1997: 83)

No Aleut dialect content questions have an overt question marker. The sentence ‘Who are you?’, for example, is *kiin ax(t)?* in Atkan, *kiin txin?* in the Eastern dialect, and *kiin tin?* in Attuan, where only the interrogative *kiin* ‘who’ marks the sentence as a question. In Atkan the interrogative is followed by a second person form of the copula *a-* ‘to be’ while in the other two dialects there is an overt second person singular pronoun that can be analyzed as *t(x)i-n* ‘DEM-2SG’ (Bergsland 1997: 57, 81, 89, 135). The marking

5 Survey of the grammars of questions in Northeast Asia

of questions in Aleut is thus typologically close to the Tungusic language Evenki, for example, although alternative questions contain a disjunction.

(53) Aleut (Atkan)

*ting asxuunulax txin satxaĥ taġaaġan aġs?*  
1SG or 2SG gill.net check.? be.?

‘Am I or are you going to check the gill net?’ (Bergsland 1997: 83, *passim*)

Copper Island Aleut presents a special case because of its strong Russian impact. Unfortunately, almost no information on interrogative constructions is available. As far as the few examples allow any conclusions, polar and content questions were probably unmarked.<sup>3</sup>

(54) Copper Island Aleut

a. *ya-ġim būd-iŝ=min’a sū-t’?*  
1SG-?REL.SG AUX-2SG=1SG.ACC take-INF

‘Will you take me with you?’

b. *aqú-ġ ti āngicikāy-usā-iŝ?*  
what-ABS.SG 2SG think-TR-2SG

‘What are you thinking about?’ (Golovko & Vakhtin 1990: 103, 104)

The auxiliary *bu(d)-*, the personal ending *-iŝ*, the infinitive marker *-t’* as well as the pronouns *ya*, *min’a* (used as a verbal person marker), and *ti* are of Russian origin.

The following short dialogue between a five year old child and her mother in Sirenik was recorded in 1985. The data show a mixed language that is comparable to Copper Island Aleut but might be more strongly based on Russian. The short dialogue includes an alternative question and an answer. There is juxtaposition of the two alternatives and there is no question marker or disjunction. Presumably, the question had a special intonation contour.

(55) ?Central Siberian Yupik

a. *mam, ya èto quuv.a=y.u, ya èto niv.a=y.u?*  
mom 1SG this pour.out=1SG 1SG this pour.into=1SG

‘Mom, shall I pour this out or shall I pour this (into something)?’

b. *ladno, quuv=a.y.*  
alright pour.out=IMP

‘All right, pour it out.’ (Vakhtin 1998: 324)

As the alternative questions appears to consist of two juxtaposed focus questions, we may surmise that focus and polar questions were marked by intonation as well.

Before focusing on Yupik as spoken in Siberia, let us have a brief look at the better-known **Central Alaskan Yupik** language to establish a reference point. Polar questions in this language are marked with a second position marker that also marks focus questions.

<sup>3</sup>Most elements in these examples are from Russian, except for those underlined, which derive from Eskaleut.

(56) Central Alaskan Yupik (General)

- a. *tekít-ùq#qaa nuk'aq?*  
arrive-3SG.IND≠Q PN.ABS.SG  
'Has Nuk'aq arrived?'
- b. *nùk'àq#qaa tekit-uq?*  
PN.ABS.SG≠Q arrive-3SG.IND  
'Has Nuk'aq arrived?' (Miyaoaka 2012: 168)

Polar questions may also be marked with final rising intonation alone. The marker *qaa* has been translated as 'right' and may also mark tag questions. It is also optionally found on the first alternative in alternative questions and combines with a disjunction.

(57) Central Alaskan Yupik (General)

- a. *enér-pa-ηqér-tuten, qaá/qáa?*  
house-big-have-2SG.IND Q  
'You have a big house, right?'
- b. *qayar-pa-li-uq#qaa wall'u# pi-cuar-mek?*  
kayak-big-make-3SG.IND≠Q or≠ thing-small-ABM.SG  
'Is he making a big kayak or a small one?' (Miyaoaka 2012: 170)

Mild questions are marked with the suffix *-ti-* translated as 'perhaps' and topic questions (polar or content) with *=mi*.

(58) Central Alaskan Yupik (General)

- a. *kai-lli-uten?*  
hungry-INF-2SG.IND  
'Are you hungry?'
- b. *kaig-tuten=qaa?*  
hungry-2SG.IND=Q  
'Are you hungry?'
- c. *u-na=mi?*  
this-EX.ABS=Q  
'How about this?'
- d. *nauwa=mi aata-ka?*  
where=Q father-ABS.1SG.SG.POSS  
'(But) where is my father?' (Miyaoaka 2012: 173, 171)

Content questions uttered in soliloquy contain an enclitic *=kiy* 'I wonder' that attaches to the initial interrogative.

- (59) Central Alaskan Yupik (General)

**qaillun=kiq tai-ga?**

how=Q          come-3SG.Q

‘How did he come over, I wonder?’ (Miyaoaka 2012: 1360)

Whether Central Siberian Yupik or Naukan Yupik share all of these question markers remains obscure from the limited and problematic publications available to me.

As in this last example, and similar to Yukaghiric languages (§5.14.2), content questions exhibit an additional interrogative mood marking on the verb that replaces declarative endings.

- (60) Central Alaskan Yupik (General)

a. *nùk’àq tekít-ùq.*

PN.ABS.SG arrive-3SG.IND

‘Nuk’aq arrived.’

b. *ki-na tekít-a-Ø?*

who-ABS.SG arrive-3Q-3SG.Q

‘Who has arrived?’ (Miyaoaka 2012: 167)

This last type of question marking shows that questions in Yupik are much more complicated than in Aleut as they combine special interrogative mood suffixes with special interrogative person endings.

Morphological question marking in Yupik involves two layers of suffixes. The first is an actual question marker and attaches to the stem (Table 5.12), followed by the second, which is an agreement marker of person and number specialized for questions. Regarding the second layer, there is a distinction between intransitive and transitive paradigms. What is more, the first layer exhibits an additional distinction into different forms that depends on person as well. There are, furthermore, some complex morphonological patterns of interactions between the stem and the two layers of suffixes that cannot be dealt with here in detail, e.g. CAY *niic+ta+yu* ‘hear+3Q+3SG.S.3SG.O.Q’ > *niitau* ‘does (s)he hear it?’ (Miyaoaka 2012: 1350). Table 5.12 lists the first layer of question marking. Generally, first and second persons are marked the same way, while third person receives another marker.

The morphosyntactic behavior of the forms is quite complex, but has only been described in sufficient detail for Central Alaskan Yupik: “The initial /c/ of the first- and second-person mood markers is fricativized to /z/ after a vowel if the subject is singular and, if the subject is non-singular, (though with some variance) after a stem that ends in a stop plus /i/.” (Miyaoaka 2012: 1352)

Intransitive interrogative agreement forms are given in Table 5.13. Apart from Sirenik, the individual affixes are very similar across the different languages.

In Naukan Yupik the form *-see* might derive from a combination of *-si(i)* with *-ŋa*. Note a parallel in the transitive paradigm below: *-see* ‘2PL.S.2SG.O’ = CAY *+ci+ŋa*. As we have just seen in Table 5.12, CAY *-ci* corresponds to Naukan Yupik *-si(i)* (~ *-jii*).

Table 5.12: Simplified inventory of interrogative mood endings in Central Alaskan Yupik (Miyaoaka 2012: 1352), Central Siberian Yupik (Jacobson 1979: 61), Naukan Yupik (Menovshchikov 1975: 240ff.), and Sirenik (Vakhtin 2000: 517)

	CAY	CSY	Naukan	Sirenik
3SG	C +ta ~ V +ya	+(t)a ~ +(g)a	-aa	-taa ~ -tsaa
3PL	id.	id.	id.	id.
3DU	id.	id.	id.	id.
1SG	+ <sub>1</sub> ci	+(t)zi	-si(i) ~ -jii	-sii ~ -tsii
1PL	id.	+ste	id.	id.
1DU	id.	id.	id.	id.
2SG	id.	+(t)zi	id.	id.
2PL	id.	+ste	id.	id.
2DU	id.	id.	id.	id.

Table 5.13: Intransitive interrogative person endings in Central Alaskan Yupik (Miyaoaka 2012: 1352), Central Siberian Yupik (St. Lawrence Island, Jacobson 1979: 61), Naukan Yupik (Menovshchikov 1975: 240), and Sirenik (Vakhtin 2000: 521)

S	CAY	CSY	Naukan	Sirenik
3SG	+∅	+∅	-∅	+a
3PL	+t	+t	-t	+i
3DU	+y	+k	-k	-
1SG	+ŋa	+ŋa	-see, -∅	+n'
1PL	+ta	+a	-ta	+ta
1DU	+nuy	+uŋ	-nuŋ	-
2SG	+t	+n	-n	+∅
2PL	+ci	+tsi	-si	+si
2DU	+tiy	+k	-tyk	-

5 Survey of the grammars of questions in Northeast Asia

Tables 5.14, 5.15, 5.16 contain transitive interrogative endings from Central Alaskan Yupik, Central Siberian Yupik, and Naukan Yupik.

Table 5.14: Transitive interrogative person endings in Central Alaskan Yupik (Miyaoka 2012: 1352); forms in parentheses are identical with the intransitive forms

A/O	3SG	3PL	3DU	1SG	1PL	1DU	2SG	2PL	2DU
3SG	+yu	+ki	+kiy	+ŋa	+kut	+kuy	+tin	+ci	+tiy
3PL	+tyu	+tki	+tkiy	+tja	+tkut	+tkuy	+tyin	+tci	+ttiyy
3DU	+yniyu	+yniki	+ynikiy	+yŋa	+ykut	+yk+kuy	+ytyin	+ytc	+yttiy
1SG	(+ŋa)	(+ŋa)	(+ŋa)				+kin <sup>4</sup>	?	?
1PL	(+ta)	(+ta)	(+ta)				?	?	?
1DU	(+nuy)	(+nuy)	(+nuy)				?	?	?
2SG	+yu	+ki	+kiy	+ŋa	+kut	+kuy			
2PL	+ci+yu	+ciki	+cikiy	+ci+ŋa	+cikut	+cikuy			
2DU	+tiyu	+tiki	+tkiy	+tiyŋa	+tiykut	+tiykuy			

Table 5.15: Transitive interrogative person endings in Central Siberian Yupik (St. Lawrence Island, Jacobson 1979: 61, 56); forms in parentheses are identical with indicative forms

A/O	3SG	3PL	3DU	1SG	1PL	1DU	2SG	2PL	2DU
3SG	+gu	+ki	+kek	+ŋa	+nkut	+nkuj	+ten	+si	+stek
3PL	+tgu	+tki	+tkek	+tja	+nkut	+nkuj	+ten	+si	+stek
3DU	+gnegu	+gneki	+gnekek	+gneŋa	+nkut	+nkuj	+ten	+si	+stek
1SG	+kun	+ngi	+gnek				+ken	(-msi)	(-mtek)
1PL	+gggu	+ki	+kek				+ken	(-msi)	(-mtek)
1DU	+gnegu	+gneki	+gnekek				+ken	(-msi)	(-mtek)
2SG	+gu	+ki	+kek	+ŋa	+kut	+kuj			
2PL	+gggu	+ki	+kek	+ŋŋa	+kut	+kuj			
2DU	+gnegu	+gneki	+gnekek	+gneŋa	+gnekut	+gnekuj			

In Sirenik, intransitive first person singular and second person plural forms are identical to the declarative endings. Paradigms for transitive verbs are almost entirely unknown. Vakhin (2000: 521) mentions *+(gy)pyŋ'/(gy)pyŋ'* '2SG.A.1SG.O', *+n'/(+n)* '3PL.A.1SG.O', *+kyn'/(+kyn)* '1SG.A.2SG.O', *+tyn'/(+tyn)* '3SG.A.2SG.O', and *+gu'/(+g)* '2/3SG/1PL.A.3SG.O'. Apart from the first, these seem to correspond to Central Alaskan Yupik *+(t)ŋa*, *+kin*, *+tin*, and *+(t/yni)yu*, respectively (Table 5.14).

<sup>4</sup>In the original table this form was given one row below, which seems to be a mistake (cf. Miyaoka 2012: 1350). Miyaoka (2012) is not sufficiently clear about the gaps marked with a question mark here. The other languages show indicative forms here that roughly correspond to CAY *-mt+yin* '1PL.2SG', *-miytin* '1DU.2SG', *-mci* '1SG.2PL', *-mtci* '1PL.2PL', *-miyci* '1DU.2PL', *-mtiy* '1SG.2DU', *-mttiy* '1PL.2DU', and *-miytiy* '1DU.2DU' (Miyaoka 2012: 1325).

Table 5.16: Tentative transitive interrogative person endings in Naukan Yupik (based on Menovshchikov 1975: 241f.)

A/O	3SG	3PL	3DU	1SG	1PL	1DU	2SG	2PL	2DU
3SG	-ûŋ	-ki	-kyk	-ŋa	-tykut	-tykuk	-tyn	-si	-tyk
3PL	-txuŋ	-tyki	-tykyk	-tŋa	-tykut	-tykuk	-tyn	-si	-tyk
3DU	?	?	?	?	?	?	?	?	?
1SG	-ku	-nka	-xka				-kyn	-msi	-mtyk
1PL	-vut	-put	-xput				-mtykyn	-msi	-mtyk
1DU	-vuk	-puk	-xpuk				?	?	?
2SG	-ûŋ	-ki	-kyk	-ŋa	-kut	-kuk			
2PL	-siûŋ	-siki	-sikyk	-see	-sikut	-sikuk			
2DU	?	?	?	?	?	?			

Some of the agreement forms in Central Alaskan Yupik (Table 5.14) are still analyzable in two different affixes (stem-A-O). According to this observation, the following suffixes can be extracted:  $-\emptyset$ - ‘3SG.A’,  $-t$ - ‘3PL.A’,  $-y(ni)$ - ‘3DU.A’,  $-\emptyset$ - ‘2SG.A’,  $-ci$ - ‘2PL.A’, and  $-tiy$ - ‘2DU.A’. These are related to, but not identical with the intransitive markers (Table 5.13). With some exceptions, these suffixes are also present in Central Alaskan Yupik and Naukan Yupik. In Central Alaskan Yupik “gaps in the paradigm are filled in with an intransitive person marker, which is extended to transitive use, without distinguishing the object number” (Miyaoaka 2012: 1350). Central Siberian Yupik on the other hand has special third person as well as second person singular object forms and employs the indicative forms as second person plural and dual object endings. In Central Siberian Yupik, the interrogative mood marker (Table 5.12) takes a form with *i* instead of *a* before the endings with subscript *i*. “The final or semi-final vowel of these endings is often lengthened (and *e* changed to *a*) if the verb is used in a ‘yes’ or ‘no’ question.” (Jacobson 1979: 61) Menovshchikov’s (1975: 242) table of interrogative forms on which Table 5.16 was based seems to be rather problematic, as it apparently shows some confusion regarding grammatical relations. My analysis usually follows the comparison with Central Alaskan (Table 5.14) and Central Siberian Yupik (Table 5.15) (see also Menovshchikov 1975: 241). For lack of data, dual A forms have usually been excluded. In some instances either intransitive or transitive verb endings may be employed with a slight change of meaning.

## (61) Central Alaskan Yupik

- a. *ca-mek*      *ner-yug-ci-t?*  
 what-ABM.SG eat-DES-2SG.Q-2SG.INTR.Q  
 ‘What (kind of food) do you want to eat?’

## 5 Survey of the grammars of questions in Northeast Asia

- b. *ca ner-yug-ci-u?*  
 what eat-DES-2SG.Q-2SG.TR.Q  
 ‘What/which (specific) food do you want to eat?’ (Miyaoka 2012: 756)

A difference between Central Alaskan Yupik and Central Siberian Yupik is that in the former they are limited to content questions while in the latter they are also encountered in polar questions.

- (62) Central Siberian Yupik (St. Lawrence Island)
- a. *negh-yug-si-n?*  
 eat-DES-2SG.Q-2SG.Q  
 ‘Do you want to eat (anything)?’
- b. *sa-meη negh-yug-si-n?*  
 what-ABM.SG eat-DES-2SG.Q-2SG.INTR.Q  
 ‘What (kind of food) do you want to eat?’ (Jacobson 1979: 60)

Table 5.17: Summary of question marking in Eskaleut.

	PQ	CQ	AQ
Aleut (Atkan)	=ii	-	asxuunulax ‘or’
Mednyj Aleut	-	-	?
CAY	V≠qaa	(see Tables 5.12, 5.13, 5.14)	(A≠qaa) + wall’u ‘or’
CSY	(see Tables 5.12, 5.13, 5.15)	id.	?
Naukan	(see Tables 5.12, 5.13, 5.16)	?id.	?
Sirenik	(see Tables 5.12, 5.13)	?id.	?

There is a marked contrast between Aleut and Yupik question marking (Table 5.17). Aleut resembles the Northeast Asian mainstream, while Yupik belongs to an area in the northern part of NEA that exhibits complex interrogative mood systems (e.g., Audova 1997). Other languages belonging to this belt are Nganasan (§5.12.2), Yukaghiric (§5.14.2), and perhaps Negidal (§5.10.2).

### 5.4.3 Interrogatives in Eskaleut

The comparison of interrogatives in Yupik languages and Sirenik is relatively straightforward (Table 5.18). The interrogative system in all four languages listed is relatively similar, but Sirenik is clearly the most aberrant. All three languages have resonances in *q~* and *n~*. Apart from CAY, there is an additional resonance in *s~*. Yupik thus has K-interrogatives. The authors also mention the form PE \**ay* ‘what did you say’ (CAY *ai*,

CSY *ay*, Naukan Yupik *ay*), but this is not a true interrogative (Fortescue et al. 2010: 62). Other interrogatives such as \**cuuq* ‘why’ or \**qanuq* ‘how’ can only be found in Inuit (Fortescue et al. 2010: 98, 310). Central Siberian Yupik interrogatives were also mostly left unexplained in Jacobson’s (2001: 49, 57, *passim*) description. Fortunately, there is a very good analysis for Central Alaskan Yupik by Miyaoka (2012: 443-461) that can be transferred to the other languages.

Similar to Tungusic, Chukotko-Kamchatkan, and Samoyedic the CSY stem *sa-* ‘(to do) what’ (CAY *ca-*) may take both nominal and verbal morphology and the forms meaning ‘why’ are derived from its verbal form. CAY has a form *ciin* ‘why’ that is a contraction of *ca-ŋan*, a third person singular causal connective mood form and cognate with CSY *sa-ŋan*. Naukan Yupik *si(i)mi* has a similar phonological development and seems to correspond directly with the third person reflexive form *sa-ŋami* in CAY and Sirenik. According to Jacobson (2001), *sa-ŋami* is a form that requires a third person singular subject while *sa-ŋameŋ* (not listed in Table 5.18) is used with third person plural forms. An interesting speciality of Yupik and Sirenik is the existence of two forms meaning ‘when’ for future and past actions that has no equivalent in NEA. The suffix *-ku* in *qa-ku-* is a future form, but Miyaoka (2012: 452f.) does not comment on the etymology of *qaŋ-vaŋ-* ‘when (PST)’ (CSY *qavŋaq*, Naukan Yupik *qamvaq*) but is of the opinion that it also derives from the stem *qa(ŋ)-*. CAY *qavci-n* ‘how many’ (CSY *qafsiin*, Naukan Yupik *qafsit*) is a plural absolutive form of the stem *qavciŋ-* (Sirenik *qafsi(ŋ-)*). All four languages above have KIN-interrogatives, although the stem really is *ki(t)-*, *ki-na* being its singular absolutive form and *kin-kut* its plural absolutive form. The selective interrogative *naliq* is apparently an unanalyzable form, and can be inflected, e.g. *nallir-put* ‘which one of us’ (cf. Miyaoka 2012: 451). The form *natən* ‘how’ is restricted to Naukan Yupik, CSY, and Sirenik (Fortescue et al. 2010: 223), while CAY has the interrogatives *qaillun ~ qaill’* and *qayu-* instead (Miyaoka 2012: 454f.). The special form *natən* is certainly connected with the stem *na-* that is ambiguous and means both ‘which’ and ‘where’. Table 5.19 compares locative interrogative paradigms in CSY and CAY. Demonstratives in Central Alaskan Yupik also have an allative ending *+vit* and variants (Miyaoka 2012: 769).

Because of a rather unsystematic presentation by Bergsland (1997: 80-83), no complete analysis of interrogatives in all the Aleut dialects can be presented here (Table 5.20). Copper Island Aleut, in addition to Aleut interrogatives, has borrowed a number of Russian interrogatives (Table 5.21). Similar to Eskimo and several other languages in NEA (Chapter 6), no other interrogative starts with the same consonant as does *kiin* ‘who’. However, the interrogative system is quite different from Yupik and Sirenik, although some similarities can be observed. The personal interrogative *kiin* ‘who’ (DU *kiinkux*, PL Eastern *kiinkun*, Atkan *kiinkus*), for example, is directly comparable. There is one major resonance in *q-*. The stem *qana-* has the same semantic scope over selective and locative meaning as does *na-* in Yupik and Sirenik. The stem *alqu-* (Attuan *aqu-*) ‘(to do/be) what, what kind/part, to be how’ is entirely absent from Eskimo, but exhibits the same ambiguity between a verbal and nominal stem as does PE \**cu-* ‘(to do) what’. The causal interrogative *alqu-l* ‘why’ likewise has a verbal basis.

5 Survey of the grammars of questions in Northeast Asia

Table 5.18: PE = Proto-Eskimo and PY-S = Proto-Yupik-Sirenik interrogatives and cognate sets according to Fortescue et al. (2010: 97, 98, 190, 223, 304, 310, 318); not all variants and dialectal forms are shown

Reconstructions	CAY	CSY	Naukan	Sirenik
PE *ki(na) 'who'	kina, PL kinkut	kina, PL kinkut	kina, PL kinkut	kina, PL kinəki
PE *kitu 'which, who'	kitu-	kitu-	kitu-	-
PE *qaku 'when.FUT'	qaku	qakun	qaku	qaku
PE *qaŋa, qaŋyaq 'when.PST'	qaŋyaq	qavŋaq	qamvaq	qaŋ
PE *qavciŋ 'how many'	qavciŋ	qafsina(t), qafsiin 'which in a row'	qafsiŋ	qafsi(γ-)
PE *na- 'where'	na-	na- 'which'	na-	na-
PE *naŋu, *na(C)uŋ 'where (is it)'	nauxu	3SG naaŋu, 3DU nakək, 3PL naaki	naa 'where'	3SG naŋəla
PE *nallir 'which'	naliq	naliq	naliq	nacəχ
PY-S *natə- 'which (part)'	natə-	natə- 'where', natu- 'which'	-	natəŋuχ 'which', natu 'where'
PY-S *natən 'how'	-	natən	natən	natən
PY-S *caŋan, *caŋami 'why'	ciin	3SG saŋan, 3SG.REFL saŋami, 1SG saŋama	si(i)mi	saŋami
PE *cu(na) '(to do) what'	ca-	sa-	suna, sa-	saŋəca 'what', sa- 'to do what'

Table 5.19: CSY (Jacobson 2001) and CAY locative interrogatives (Miyaoka 2012)

Meaning	CSY	CAY
where (-LOC)	na-ni	na-ni
whence (-ABL)	na-ken	na-ken
whither (-ALL)	na-vek	na-mun

Table 5.20: Atkan Aleut interrogatives according to Bergsland (1997: 80ff.)

Meaning	Form
who	kiin, DU kiin-kux, PL kiin-kus
what (-ABS.SG)	alqu-ġ
to do what (-V-)	alqu-sa-
which	qanan, PL qana-kus
where	qanang
why (-CVB)	alqu-l
how	qanamat-, qanamasix
how many/much	qanaang
when	qanayaam

Table 5.21: Interrogatives in Copper Island (Mednyj) Aleut (Sekerina 1994: 26) in comparison with Attuan Aleut (Bergsland 1997: 80ff.) and Russian (§5.5.3.3)

	Mednyj Aleut	Attuan Aleut	Russian
who	kiin	kiin	-
what (-ABS.SG)	aqu-ġ ~ aqo-ġ	aqu-ġ	-
how	aqu-ta-l	aqu-ta-	-
why	aqo-li, aqu-ma	aqu-	-
where from	qanaaga	qanaax	-
how many/much	qanan	qanang	-
when	ka(g)da	-	kokda/когда
which	kakuy	-	kakoj/какой
which	katorəye	-	kotoguj/который
where	kuda	-	kuda/куда

## 5.5 Indo-European

### 5.5.1 Classification of Indo-European

According to *Glottolog* (Hammarström et al. 2016), Indo-European encompasses 583 languages. Similar to §5.9 on Trans-Himalayan, this section can only deal with a minor part of the whole Indo-European family. The exact internal phylogenetic structure of the family is not absolutely clear (see §2.5), but one may roughly distinguish 10 different branches as well as a couple of unaffiliated and sparsely attested languages that are excluded here (Fortson 2010: 10): 1. Albanian, 2. †Anatolian, 3. Armenian, 4. Balto-Slavic (Baltic, Slavic), 5. Celtic, 6. Germanic, 7. Greek, 8. Indo-Iranian (Indo-Aryan, Iranian, and perhaps Nuristani), 9. Italic, and 10. †Tocharian. Only West Germanic (German dialects, Yiddish, English), East Slavic (Russian, Ukrainian), East Iranian (Sogdian, Khotanese, Tumshuqese, Sarikoli), and Tocharian (Tocharian A, B, and perhaps C) have representatives in NEA. For the mixed Persian-Uyghur language Eynu (*àinũ* 艾努), spoken in the southeast of the Tarim basin, see §5.11. Taimyr Pidgin Russian (or Govorka) and Chinese Pidgin Russian (sometimes called Kyakhta Pidgin) will be included in this chapter, but the mixed Russian-Aleut language Mednyj Aleut has been treated in §5.4 on Eskaleut.

### 5.5.2 Question marking in Indo-European

#### 5.5.2.1 Question marking in Proto-Indo-European

PIE presumably had interrogatives in initial position, optionally preceded by a topicalized element (Fortson 2006: 232). Questions in PIE were probably primarily marked with a special intonation contour (Delbrück 1900: 259–288; Lehmann 1974: 101f., 120–123, 179f.; Hackstein 2013: 99), although word order change is attested in several Indo-European branches (Hackstein 2013: 102). Some old Indo-European languages had sentence-initial or second position clitics (e.g., Gothic *an*, =*u*, Braune & Heidermanns 2004). However, the markers in individual branches are not cognates of each other, which is why no such marker can be reconstructed.

#### 5.5.2.2 Question marking in Germanic

Modern Germanic languages generally have verb-initial word order for marking polar questions (63b). In declarative sentences the verb usually takes second position (63a). Consider the following constructed German examples as well as their English translation. In addition, the German polar question has a rising intonation as opposed to the falling intonation in the declarative sentence.

(63) German<sup>5</sup>

- a. *Peter hat einen Hund gekauft.*  
 PN has a.ACC.M dog bought  
 ‘Peter bought a dog.’
- b. *Hat Peter einen Hund gekauft?*  
 has PN a.ACC.M dog bought  
 ‘Did Peter buy a dog?’

If no other auxiliary is present, English requires the addition of the auxiliary *to do*. As further explained in Chapter 4, the cross-linguistically untypical phenomenon of word order for question marking (Dryer 2013j) originated in the loss of a second position clitic such as Gothic =*u*. Such clitics usually attach to the verb in polar questions and to focused elements in focus questions. When the question marker was lost, the verb-initial word order took over its function (e.g., Miestamo 2011). Plautdiitsch likewise preserves the verb-initial word order.

## (64) Altai Low German

*väitst dyy va<sup>u</sup>t diinə fryy feelə deed?*  
 know.PRS.2SG 2SG what 2SG.GEN.F wife miss.INF do.PRS.3SG

‘Do you know what problem your wife has?’<sup>6</sup> (Jedig 2014: 170)

An exception among Germanic languages is Yiddish, which has borrowed the Polish, Ukrainian, or Belorussian initial question marker, which will be discussed in §5.5.3.3. Nevertheless, there is still a word order change as well as final rising intonation as opposed to the falling declarative intonation.

## (65) Yiddish

- a. *mojše hot gekojft a hunt.*  
 PN has bought a dog  
 ‘Moses bought a dog.’
- b. *ci hot mojše gekojft a hunt?*  
 Q has PN bought a dog  
 ‘Did Moses buy a dog?’ (Sadock & Zwicky 1985: 181)

The German examples in (63) above that were constructed on the basis of these Yiddish examples exhibit in addition a slightly different word order in that the participle stands sentence-finally. The Yiddish word order is not usually found in German and has an archaic flair to it. The initial question marker is optional.

<sup>5</sup>The glossing in this chapter is somewhat simplified and relies on the relatively close relationship of English with other languages.

<sup>6</sup>Cf. non-standard German (constructed) *Weißt du, was deiner Frau fehlen tut?*

- (66) Yiddish  
**bist**            *du meshuge?*  
 be.PRS.IND.2SG 2SG crazy  
 ‘Are you crazy?’ (Jacobs et al. 1994: 408)

- (67) German  
**Bist**            *du verrückt?*  
 be.PRS.IND.2SG 2SG crazy  
 ‘Are you crazy?’<sup>7</sup>

Focus questions in German and English have the same structure as polar questions but contain an additional intonational nucleus on the focused element. English may also make use of a cleft, e.g. *Is it to school that you are going?*

- (68) German  
*Gehst du zur \_\_\_\_\_ Schule ?*  
 go.PRS.2SG 2SG to.the.F.SG.DAT school  
 ‘Are you going to school?’

I am unaware of any descriptions of focus questions in Yiddish or Plautdiitsch but it is probable that they have a pattern similar to German.

Content questions in German, Plautdiitsch and Yiddish do not have a special marking but do have sentence-initial interrogatives. They may be preceded by a conjunction such as German *und* ‘and’.

- (69) German (Colloquial)
- a. **Wo** *is(t) der Mensch?*  
 where is the.M.NOM person  
 ‘Where is the person?’
  - b. *Na, und warum?*  
 well and why  
 ‘Well, why then?’

- (70) Yiddish  
**vu** *iz der mentsh?*  
 where is the.M.NOM person  
 ‘Where is the person?’ (Jacobs et al. 1994: 408)

- (71) Altai Low German  
*na, on vu:ʀqm?*  
 well and why  
 ‘Well, why then?’ (Jedig 2014: 170)

---

<sup>7</sup>Colloquially, German also has the adjective *meschugge* ‘crazy’.

Yiddish has an optional marker =*zhe* that may attach to interrogatives and seems to intensify the sentence (it was translated as ‘on earth’) (Jacobs et al. 1994: 413). It is of West Slavic origin, e.g. Czech =*že* (Sussex & Cubberley 2006: 317). English differs somewhat from these three languages in that content questions usually require an auxiliary or *to do* to follow the interrogative.

(72) English

- a. **Where is the person (going)?**
- b. **Where did he go?**

Alternative questions in German combine usual polar question marking (verb first word order and intonation) with the disjunction *oder* [-v] ‘or’. Negative alternative questions have the standard negator *nicht* that in colloquial speech often takes the form *nich*. English has a similar polar question marking in combination with *or*.

(73) German

- a. **Magst du Tee oder Kaffee?**  
want.PRS.IND.2SG 2SG tea or coffee  
 ‘Do you want tea or coffee?’
- b. **Magst du Tee oder nicht?**  
want.PRS.IND.2SG 2SG tea or not  
 ‘Do you want tea or not?’

Altai Low German has verb first word order in combination with *öuda nich* (= German *oder nich(t)*) for negative alternative questions and probably *öuda* (= German *oder* [-v]) for plain alternative questions (Nieuweboer 1999: 177). Yiddish also has the verb-in initial position but exhibits alternation between the use of *oder* ‘or’ and *ci*, which has been influenced by Slavic (Jacobs 2005: 205).

German has two different constructions in which the disjunction *oder* takes sentence-final position and acts as a question marker. In the first case it is accompanied by a longish and level intonation, in the second with a sharp rise in intonation. The former is a fully elliptic alternative question and the latter a tag question.

(74) German

- a. **Magst du Kaffee oder ...?**  
want.PRS.IND.2SG 2SG coffee or  
 ‘Do you want coffee or not?’
- b. **Du magst Kaffee, oder?**  
2SG want.PRS.IND.2SG coffee or  
 ‘You want coffee, right?’

In the latter type the tag may also take the negative form *oder* (*etwa*) *nich(t)* with an optional emphatic marker. German has several more tags such as *ja* ‘yes’, *richtig* ‘right’ and *nicht(t)* ‘not’, or dialectal forms such as *wa(t)* (Standard German *was* ‘what’) and the synchronically unanalyzable form *ge(lle)*. In all cases a tag seems to be accompanied with a sharp rising intonation contour. English has related tags such as *right* but is best-known for its tags whose polarity depends on the preceding declarative, e.g. *is it* vs. *isn’t it*, *do you* vs. *don’t you* etc. Plautdiitsch has the tag question markers *jau?* ~ *jo?* ‘yes’ (German *ja*), *nee?* (German *nein*, colloquially *ne(e)*), *öuda* ‘or’ (German *oder*), and *es nich zöu?* ‘isn’t it’ (German *is(t) es nich(t) so?*) attached to the end of a declarative sentence (Nieuweboer 1999: passim).

Indirect polar questions require a special marker, English *if* or *whether*, Plautdiitsch *ous*, or German *ob*. Interestingly, English *whether* historically derives from the PIE interrogative *\*k<sup>w</sup>óteros* ‘which of two’ (see §5.5.3.1), German *ob* and English *if* show connections with conditionals, but the etymology of Plautdiitsch *ous* is not perfectly clear. Yiddish has adopted the use of *ci* in indirect questions from Slavic.

- (75) Altai Low German  
*ous ät’ siinə fryy ka<sup>u</sup>n heilə*  
 if 1SG 3SG.GEN.F wife can.?IND.PRS.1SG cure.INF  
 ‘whether I can cure his wife’ (Jedig 2014: 170)<sup>8</sup>

In German an embedded polar question can also stand on its own to form a question usually addressed to oneself and roughly meaning ‘I wonder’.

- (76) German  
*(Ich frage mich,) ob es wohl regnen wird.*  
 1SG ask.PRS.IND.1SG 1SG.ACC if it perhaps rain.INF will.PRS.IND.3SG  
 ‘I wonder whether it will rain.’

Indirect content questions have an interrogative instead of the mentioned indirect question marker found in polar, focus, or alternative questions. Indirect content questions in German and English have the interrogative in initial position, but have a different word order from plain content questions. In German, verbs are strictly final in both types of indirect questions.

- (77) German  
*(Ich frage mich,) wer das wohl ist?*  
 1SG ask.PRS.IND.1SG 1SG.ACC who that perhaps is  
 ‘I wonder who that will be.’

Indirect content questions may also be used on their own for self questions. Both types of indirect questions are almost obligatorily accompanied with the modal marker *wohl* (cognate with English *well*).

<sup>8</sup>The word order of Plautdiitsch is also possible in German but sounds very archaic. A German equivalent would be something like the following: *ob ich seine Frau heilen kann*.

## 5.5.2.3 Question marking in Slavic

Most Slavic languages have a second position polar question clitic =*li* (Sussex & Cubberley 2006: 359). In Russian it is found especially in the written language. It also marks focus questions, in which case the focused element instead of the verb has to take sentence-initial position.

(78) Russian

- a. *otvétil=li*                    *studént na vsé voprósy?*  
 answer.M.SG.PST=Q student to all questions  
 ‘Did the student answer all the questions?’
- b. *studént=li otvétil*                    *na vsé voprósy?*  
 student=Q answer.M.SG.PST to all questions  
 ‘Was it the student who answered all the questions?’ (Sussex & Cubberley 2006: 359)

Only some languages lack the clitic but have a sentence-initial particle instead, including Ukrainian *čy/чи*, Belorussian *ci/ці*, and Polish *czy*, which has been borrowed by Yiddish.

(79) Ukrainian

- čy zdoróvyj tý?*  
 Q healthy 2SG  
 ‘Are you well?’ (Sussex & Cubberley 2006: 359)

In Ukrainian there is a sharp rise at the end of the sentence in polar questions, or over the focused element in focus questions (Shevelov 1993: 978). But this is less pronounced if the question is already marked with *cy*. In focus questions it is also possible to move the focused element into sentence-initial position.

(80) Ukrainian

- a. *cy ty tam buv?*  
 Q 2SG there were  
 ‘Were you there?’
- b. *ty buv tam?*  
 2SG were there  
 ‘Were you there?’
- c. *buv ty tam?*  
 were 2SG there  
 ‘Were you there?’
- d. *tam ty buv?*  
 there 2SG were  
 ‘Were you *there*?’ (Shevelov 1993: 978)

## 5 Survey of the grammars of questions in Northeast Asia

Interrogatives are usually fronted in both languages but not necessarily so in Russian. Content questions remain unmarked in both languages.

- (81) Russian  
*čto eto?*  
what this  
'What is this?' (Comrie 1984: 23)
- (82) Ukrainian  
*de ty був?*  
where 2SG were  
'Where were you?' (Shevelov 1993: 979)

In Russian the intonational nucleus can most often be found on the interrogative itself (Comrie 1984: 24).

Alternative questions in Russian and Ukrainian are quite different from each other. Ukrainian uses the polar question marker in between the two alternatives. Given its syntactic behavior in polar questions, it may perhaps be said to attach to the beginning of the second alternative. In Russian, on the other hand, there is a disjunction *ili* 'or'.

- (83) Ukrainian  
*ty був у театрі, чи в музеї?*  
2SG were at theater Q at museum  
'Were you at the theater or in the museum?' (Shevelov 1993: 978)
- (84) Russian  
*вы хотите чай, или кофе?*  
2PL want tea or coffee  
'Do you want tea or coffee?' (Comrie 1984: 23)

In Russian the first alternative takes neutral polar question intonation with a sharp rise and immediate less sharp fall on *čaj*, the second alternative has falling intonation similar to an interrogative in content questions. Ukrainian negative alternative questions take *чи ну/чи ні* 'Q NEG' (Pugh & Press 1999: 285). In Russian the form *или нет/или нет* is used (elicited in June 2017).

Russian uses question tags less frequently than English or German. But one possibility is to attach *не правда ли* 'NEG truth Q' to a declarative sentence (Comrie 1984: 32). Russian furthermore has the question marker *разве* that may stand sentence-initially and less frequently sentence-finally or be adjacent to the focus. Comrie (1984: 21f.) describes the use of *разве* as follows: "the questioner had a certain prior expectation; some piece of new information leads the questioner to believe that his prior expectation may be wrong; therefore he asks the appropriate general question with *разве*".

- (85) Russian  
*razve ty uezžaeš?*  
 Q 2SG leave  
 ‘Are you leaving?’ (Comrie 1984: 22)

Intonation is again similar to plain polar questions. Topic questions are introduced with the conjunction *a* ‘and’, e.g. *a viktor?* ‘what about Victor?’ (Comrie 1984: 27f).

Questions in *Chinese Pidgin Russian* seem to be generally unmarked. Interrogatives remain in situ. Interestingly, the Chinese A-not-A pattern is also possible, e.g. *pravda ne pravda?* ‘true NEG true’ (Shapiro 2010: 39), cf. Mandarin *duì-bu-duì?*

- (86) Chinese Pidgin Russian
- a. *za vashe zh'onusheki mes'aca posidi esa?*  
 TOP 2SG.GEN wife.?PL together sit REL  
 ‘Do you sit together with your wives?’<sup>9</sup>
  - b. *ni-dy šýma múr.mur?*  
 2SG-GEN what say  
 ‘What are you saying?’ (Shapiro 2010: 37, 15)<sup>10</sup>

In *Taimyr Pidgin Russian* polar questions may also be unmarked. But perhaps there is a special intonation contour in both pidgins.

- (87) Taimyr Pidgin Russian  
*tebja urusé-to jest?*  
 2SG rifle-HL EX  
 ‘Do you have a rifle?’ (Stern 2005: 312)

The suffix *-to* glossed as “highlighter” usually has a discourse function and is of northern Russian origin (see Stern 2005: 309; 2012: 439). Content questions are unmarked and interrogatives are either sentence-initial or preverbal (Stern 2012: 508).

- (88) Taimyr Pidgin  
*čego tebja nado menja čum mesto?*  
 what 2SG DEON 1SG tent place  
 ‘What do you want in my tent?’ (Stern 2012: 361)

The polyfunctional case marker *mesto*, from a noun meaning ‘place’, is an innovation of the pidgin (Stern 2012: 360–382). There is an instance of an open alternative question that combines a disjunction, an interrogative, and a (polar) question marker in that order.

<sup>9</sup>The topic marker *za* stems from Mongolic (Shapiro 2010: 35).

<sup>10</sup>Cf. Mandarin *nǐ-de* 你的 ‘2SG-GEN’, *shénme* 什么 ‘what’.

(89) Taimyr Pidgin Russian

*ty govorit, mama əmədja ne puskat budem, ili čego=li?*  
 2SG say.3SG mother hither NEG let.INF AUX.FUT or what=Q

‘Are you saying that we should not let mother come here or what?’ (Stern 2005: 310)

The use of a disjunction may be traced to Russian influence, but the presence of the Russian polar question marker after an interrogative might be local influence. Compare Nganasan open alternative questions in which the second part takes a similar form, e.g. *maa-ŋu-* ‘what-AOR.Q-’ (§5.12.2). However, similar phenomena are also known from Russian. For example, the following sentence was recorded in the Allaikhovsky district in the north of the Sakha Republic: *A, zdec’ čto=li?* ‘So here or what?’ (Krasovitsky 2004)

5.5.2.4 Question marking in Iranian

Polar questions in the extinct Iranian language Saka are unmarked except for, perhaps, intonation. Negative alternative questions are marked with a disjunction *aa* that later developed into *o*, followed by the negator *ne*. Content questions seem to have remained unmarked and interrogatives were fronted (Emmerick 2009: 402). Optionally a “discourse initiator” (e.g., *tta* ‘thus, so’) could precede an interrogative (Emmerick 2009: 403), which is typologically similar to initial Mandarin *nà* 那 ‘that’ or English *so*.

(90) Saka (Tumshuqese)

*uvāsānu samvāru paitryai pātanāya?*  
 laywoman.GEN.PL regulations.ACC.SG can.2SG.PRS.IND keep.INF

‘Can you keep the regulations of the laywomen?’ (Emmerick 1985: 10f., 14)

(91) Saka (Khotanese)

*cuude bre mā?*  
 why weep.2SG.IND

‘Why are you weeping?’ (Emmerick 2009: 402)

**Sogdian** has an optional sentence-initial polar question marker *(ə)ču-t(i)* ‘what-COMP’ (Yoshida 2009: 316f.). Negative alternative question seem to have the same marker at the beginning of the whole sentence in combination with a marker *kataar(-əti)* ‘which(-COMP)’ between the two alternatives. Sogdian thus has two question markers that derive from interrogatives. Sogdian *kataar*, like English *whether*, derives from PIE \**k<sup>w</sup>ótero-* ‘which of two’ (§5.5.3). Content questions have no special marking. Rhetorical questions have in addition a marker *pnuukar*. Interrogatives remain *in situ*.

(92) Sogdian

a. *ə.ču.ti pnuukar tawa waanoo nee patyoošti?*  
 Q Q.RHET by.you thus NEG heard.PRET

‘Have you never heard this?’

b. *ču.ti xa zaakt taδeeδ ay atant kataar.əti nee?*

Q the children there came Q NEG

‘Have the children arrived at your place or not?’

c. *tay u peernamstar ču əktya kθaare?*

2SG before what deed do.PRET.2SG

‘What was it that you did before?’ (Yoshida 2009: 317)

There are initial question markers in Persian (*aayaa*) and Tajik (*oyo*) as well, but these show no connection with interrogatives (Windfuhr & Perry 2009: 438). Tajik may also employ the Uzbek sentence-final marker =*mi* (Windfuhr & Perry 2009: 481, §5.11.2).

For Sarikoli there is a relatively old description by Shaw (1876: 29). According to him, polar questions in Sarikoli have a sentence-final marker *â*, while content questions remain unmarked. This marker, according to Gao Erqiang (1985), has the form *o* and has been reanalyzed here as an enclitic.

(93) Sarikoli

*bofa=af tag wand=o?*

PN=2PL actually see.PST=Q

‘Have (you) actually seen Bosha?’ (Gao Erqiang 1985: 62)

The same marker appears twice in alternative questions. The following example contains in addition an element *naji* in between the two alternatives that was glossed as a negator but seems to have the function of a disjunction here.

(94) Sarikoli

*maf tuχw χ or-an=o naji wi budo χ or-an=o?*

1PL chicken eat-1PL=Q NEG that beef eat-1PL=Q

‘Do we eat chicken or do we eat that beef?’ (Gao Erqiang 1985: 65)

According to Gao Erqiang (1985: 118), Wakhi has a disjunction *jo* and question marking on the first alternative only (=a), which is a construction similar to surrounding languages such as Uyghur (§5.11.2). In Sarikoli there are several tag question markers that contain the same question marker, e.g. *na sou-d=o?* ‘NEG be.possible-3SG=Q’ (Gao Erqiang 1985: 90).

(95) Sarikoli

*tudzik ziv ati wazon-d, rust=o?*

PN language ? know.N.PST-3SG true=Q

‘(She) knows Sarikoli, right?’ (Gao Erqiang 1985: 89)

The question marker seems to be connected with Burushaski, e.g. *bás=a?* ‘Is it enough?’, and several surrounding languages (Yoshioka 2012: 190). Consider an example from the Dardic language Palula spoken in the extreme north of Pakistan where the marker, depending on the dialect, takes the form =*aa* or =*ee*.

## 5 Survey of the grammars of questions in Northeast Asia

- (96) Palula (Dardic, Indo-European)  
*búd-u=ee?*  
understand.PFV-MSG=Q  
‘Did you understand?’ (Liljegren 2016: 403)

Content questions in Sarikoli are unmarked and interrogatives seem to remain *in situ*.

- (97) Sarikoli  
*muu vrud tar ko?*  
1SG.GEN brother EX where  
‘Where is my brother?’ (Gao Erqiang 1985: 86)

### 5.5.2.5 Question marking in Tocharian

Tocharian has unmarked polar questions (98), but perhaps had a special intonation contour that cannot be reconstructed.

- (98) Tocharian B  
*ate kampāl yamašasta?*  
away coat.ACC do.PRET.2SG  
‘Have you put (your) coat away?’ (Hackstein 2013: 110)

In Tocharian A there are two optional question markers, second position *=te* and sentence-final *aśśi* that may be found together in one sentence.

- (99) Tocharian A  
*ynālek=te lo kāl̄k aśśi?*  
elsewhere=Q away go.PRT.3SG Q  
‘Has (s)he gone somewhere else?’ (Hackstein 2013: 111)

According to Hackstein (2004: 175) *aśśi* derives from PIE *\*h<sub>2</sub>et + \*k<sup>w</sup>ih<sub>1</sub>* (cf. Latin *atquī*), of which the latter part is an instrumental form of an interrogative that is also the source of, for example, Polish *czy* (cf. Latin *quī* ‘how’). The first part *\*h<sub>2</sub>et*, or *\*h<sub>a</sub>et* with one of the laryngeals *h<sub>2</sub>* or *h<sub>4</sub>* according to Mallory & Adams (2006: 289ff.), is a preposition meaning ‘away, beyond’ (e.g., Tocharian B *at(e)* ‘away’). Hackstein assumes that *aśśi* started out as a question tag similar to German *wie* ‘how’ or *was* ‘what’. Content questions may be unmarked in both Tocharian B (e.g., Adams 2013: 157) and Tocharian A, e.g. *kus tām?* ‘Who is that?’ (Carling 2009: 156). The particle *aśśi* can also be encountered in content questions and sometimes fuses with interrogatives, e.g. *tā, tāśśi* ‘where’ (Sieg & Siegling 1931: 182). In some instances the last consonant of the interrogative is lengthened, e.g. *kus, kuss aśśi* ‘who, what’ (Sieg & Siegling 1931: 190). In Tocharian A some interrogatives are also often followed by *pat (nu)*, the exact function of which remains unclear to me, e.g. *kus pat nu* ‘or what now’ (Carling 2009: 156). In alternative questions one marker appears on each alternative in Tocharian B.

- (100) Tocharian B  
*pañākte=wat yopsa, nānde=wat?*  
 PN.NOM=Q enter.PRET.3SG PN.NOM=Q  
 ‘Has Buddha or Nanda (just) entered?’ (Hackstein 2013: 110)

Interestingly, Hackstein (2013: 111) also has an example of a negative alternative question that has a disjunction, Tocharian B *epe mā?* ‘or NEG’. The same disjunction can optionally also be found in plain alternative questions (Hackstein 2013: 111). Tocharian A has negative alternative questions with the marker *=te* used once on each alternative. In the second alternative it attaches to the negator, *mā=te*. In one such example there is an additional element *na* in the second alternative, glossed as a question marker by Hackstein.

- (101) Tocharian A  
*cāmpāl=te nasan cesām wrasaśsi waste māskatsi,*  
 be.able.GER2.NOM=Q COP.PRS.1SG this.GEN.PL being.GEN.PL refuge be.INF  
*mā=te cāmpāl (na) sam?*  
 NEG=Q be.able.GER2.NOM (Q) COP.PRS.1SG  
 ‘Am I able to provide refuge to the beings or am I not able?’ (Hackstein 2013: 113)

The double marking strategy optionally combines with the disjunction *epe* ‘or’.

- (102) Tocharian A  
*mā=(t)e nātāk cam brā(maṃ) e(pe) mā=(t)e was?*  
 NEG=Q master this.ACC.SG PN or NEG=Q 1PL.ACC  
 ‘Will the master not keep this Brahman or will he not keep us?’ (Hackstein 2013: 113)

### 5.5.2.6 Summary

Question marking in Indo-European is very different from the majority of languages in NEA. As expected for a family with such a long history, the marking of questions varies strongly from language to language. Even within the relatively shallow Slavic branch there are marked differences. Generally there is a tendency for initial particles or second position clitics and disjunctions. To the best of my knowledge, almost all Indo-European languages included here have unmarked content questions. Interestingly, at least four languages (German *wie*, *was*, Ukrainian *čy* (hence Yiddish *ci*), Sogdian (*ə*)*ču-t(i)*, *kataar(-əti)*, and Tocharian A *aśši*) show a development from interrogative to polar question marker and/or question tag, which is quite unusual for NEA (but see §5.12.2 on Selkup). This is also known from other Indo-European languages, such as Sanskrit *kad* ‘what’ (Hackstein 2013: 100), Bengali *ki* ‘what’ (Thompson 2012, see §4.2.1), or Palula *ga* ‘what’ (Liljegren 2016, §4.2.1, §4.2.3).

Table 5.22: Summary of question marking in Indo-European languages

	PQ	CQ	AQ
English	#V + (to do)	(to do)	#V + (to do) + or
German	#V	-	#V + oder 'or'
Plautdiitsch	#V	-	?#V + öuda 'or'
Yiddish	#ci V	-	#V + odər 'or', #V + X ci Y
Ukrainian	#čy	-	X čy Y
Russian	-, #V=li	-	ili 'or'
	-	-	?ili 'or' + =li
Taimyr Pidgin			
Russian			
CPR	-	-	?
Sogdian	#(ə)ču.t(i)	-	#(ə)ču.t(i) + X kataar(-əti) Y
Khotanese	-	-	aa (> o) 'or'
Tumshuqese	-	-	o 'or'
Sarikoli	=o#	-	2x=o# + ?naji 'NEG'
Tocharian A	#A=te, ?aśsi#	-, ?aśsi	2x =te (+ epe 'or')
Tocharian B	-	-	2x =wat, epe 'or'

### 5.5.3 Interrogatives in Indo-European

#### 5.5.3.1 Interrogatives in Proto-Indo-European

A somewhat outdated, but nevertheless useful, typological classification of Indo-European is in so-called *centum* and *satəm* languages. The designation follows the Latin and Avestan words for 'hundred', respectively, that represent the two types. The two groups are divided by their reflexes of Proto-Indo-European velar, palatal and labiovelar consonants. In *centum* languages the palatals and in *satəm* languages the labiovelars became plain velars (Table 5.23). PIE *\*k<sub>m</sub>tom* 'hundred' starts with a palatal and thus remained a palatal in Iranian (and later changed to *s* in Avestan) but became a plain velar written as <c> in Latin (cf. Fortson 2010: 146). The languages in this study belong to both the *centum* (Germanic, Tocharian) and *satəm* (Iranian, Slavic) types.

This division is important for the purposes of this study, because PIE interrogatives usually began with the labiovelar *\*k<sup>w</sup>* that was preserved in Germanic and Tocharian, but changed to plain velars in (Indo-)Iranian and (Balto-)Slavic. In Tocharian the labiovelars were later mostly lost in favor of plain velars. However, in some instances they show reflexes, e.g. Tocharian A *kus*, B *k<sub>u</sub>se* 'who'. In Germanic *\*k<sup>w</sup>* regularly changed to *\*h<sup>w</sup>*, e.g. Old English *hwā* or Old High German (*h*)*wer* 'who'. German later entirely lost the initial consonant, e.g. German *wer* 'who'. In English the development is more complicated. The modern spelling preserves the original <hw> with metathesis as <wh>, but the pronunciation varies between /h/ (*who*) and /w/ (*what*). In Slavic as well as Iranian, plain velars were palatalized before front vowels such as *i* but otherwise remained stable,

e.g. Russian *kto* ‘who’ (PIE  $*k^w o-$  +  $*tod$ ) but *čto* ‘what’ (PIE  $*k^w i/e-$  +  $*tod$ ), or Avestan *kas(a)-* etc. ‘who’ (PIE  $*k^w ó-s$ ) but *ciš* ‘who’ (PIE  $*k^w í-s$ ) (e.g., Fortson 2010: 231f., 421).

Table 5.23: Developments of PIE velars (Fortson 2010: 58)

centum <	PIE	> satəm
*k	*k̥	*k̥
	*k	*k
*k <sup>w</sup>	*k <sup>w</sup>	

PIE interrogatives can be reconstructed as  $*k^w o-$ ,  $*k^w e-$ ,  $*k^w i-$ , and  $*k^w u-$  (e.g., Cysouw & Hackstein 2011), but a controversy concerns the status of the last of them. Dunkel (2014: 436–441) argues that it must be reconstructed as  $*kú-$  ‘where’ and thus does not belong to the group of interrogatives starting with  $*k^w-$ . According to him, the forms beginning with  $*k^w-$  were actually derived from  $*kú$  in the first place (Dunkel 2014: 436). In his analysis,  $*k^w i-$  and  $*k^w e-$  are combinations of  $*kú-$  with anaphoric stems  $*i-$  and  $*e-$ , while the formation of  $*k^w ó-$  is not solved entirely (see Dunkel 2014: 478). Whether this hypothesis is correct cannot be decided here, but it seems possible. If it is accurate, interrogatives in Indo-European are ultimately based on locative interrogatives as in German (see §5.5.3.2).

Table 5.24 gives the list of interrogatives in PIE as reconstructed by Mallory & Adams (2006). A more extensive discussion of interrogatives can be found in Dunkel (2014: 453–479). It cannot be given here in its entirety, although some of his reconstructions have been integrated into Table 5.24. In some cases there is a corresponding demonstrative with the same endings, e.g.  $*to-deh_a$  ‘then’,  $*to-r$  ‘there’,  $*to-ti$  ‘so much/many’ etc. According to the authors,  $*k^w om$  ‘when’ is a masculine accusative form of  $*k^w os$  ‘who’, which seems extremely unlikely from a semantic point of view.

Proto-Indo-European thus had K-interrogatives but no KIN-interrogative (because of a missing nasal). Perhaps, the relative  $*yo-$  goes back to an interrogative as well but is not attested as such (Mallory & Adams 2006: 421). There may have been one interrogative stem that did not start with  $*k^w-$  but  $*m-$ . It has been reconstructed as  $*me/o-$  ‘who, which’ (Mallory & Adams 2006: 421; Cysouw & Hackstein 2011; Dunkel 2014: 518–523), e.g. Tocharian A *mānt* ‘how’.

### 5.5.3.2 Interrogatives in Germanic

Table 5.25 gives the diachrony of several Germanic interrogatives and their modern German and English cognates. For some additional discussion see also Dunkel (2014). PIE  $*k^w otero-$  ‘which of two’ has lost its interrogative meaning in German *weder* ‘neither’ and in English *whether*, used for indirect polar, focus, and alternative questions.

German, Yiddish, and Plautdiitsch share a single resonance in  $v-$ , as German <w> is pronounced as [v] as well (Table 5.26). As mentioned before, English has a variation between  $w-$  and two forms starting with  $h-$ . Altai Low German *vənäiv* ‘when’ is closer

5 Survey of the grammars of questions in Northeast Asia

Table 5.24: Selected PIE interrogatives with some cognates according to Malory & Adams (2006: 419f.); extended with the help of Dunkel (2014: 436-441, 453-479); accents partly removed

PIE form	PIE meaning	Selected Cognates
*k <sup>w</sup> ós	who NOM.SG.M	NE who, Grk toû, Skt kas, Got lvas
*k <sup>w</sup> ís	who NOM.SG.AN	Lat quis, Grk tis, Av ciš
*k <sup>w</sup> ód	what NOM/ACC.SG.N	Lat quod, NE what, Skt kad
*k <sup>w</sup> íd	what NOM/ACC.SG.N	Lat quid
*k <sup>w</sup> óm	?when ACC.SG.M	Got hvan, OCS ko-gda, (Lat cum)
*k <sup>w</sup> ih <sub>1</sub>	how INST	Lat quī, AE hwī, Polish czy
*k <sup>w</sup> óteros	which (of two)	Grk poteros, Skt katarā-, OCS koteryjǐ
*k <sup>w</sup> odéh <sub>a</sub>	when	Skt kadā
*k <sup>w</sup> or	where	Lat quōr, OHG hwār, Skt karhi
*k <sup>w</sup> óti	how much/many	Lat quot, Grk posos, Skt kati
*k <sup>w</sup> éti	how much/many	Av čaiti, Breton pet der ‘how many days’
*k <sup>w</sup> eh <sub>a</sub> k-	what kind of	[North West] OCS kakŭ, Lithuanian kok(i)s
*k <sup>w</sup> oli	?how much	only OCS kolikŭ ‘how large’, kolǐ ‘how much’
*k <sup>w</sup> u(ú), ?*kú	where	Lat ubi, Grk pu-, Skt kŭ
*k <sup>w</sup> u-d <sup>h</sup> e	where	OCS kŭde, OAv kudā, Skt kuha

to Dutch *wanneer* than to German *wann*. Similar to Dutch *waar* and English *where* it also retains a reflex of a final *r* in *vuue*, while German only preserves an older form *wor-* in derived forms. But the form *vou-* directly corresponds to German *wo-*. Also compare German *worauf*, Dutch *waarop*, and Plautdiitsch *vourōp* ‘on what’ as well as German *was*, Dutch *wat*, and Plautdiitsch *vaut* ‘what’. Yiddish *far vos* has direct correspondences in German *für was* and English *what for*. This is a common European formation, e.g. Italian *perché* ‘why’ (cf. *per* ‘for’, *che* ‘what’).

German exhibits an interesting congruence of the two forms *was* ‘what’ and *wie* ‘how’ that in certain circumstances are mutually exchangeable.

(103) German

- a. **Wie/Was ist dein Name?**  
how/what is your.M.SG name  
‘What is your name?’
- b. **Bist du fertig oder wie/was?**  
are 2SG ready or how/what  
‘Are you ready or what?’
- c. **Du bist der Neue, wie/was?**  
you are the.M.SG new.one how/what  
‘You are the new one, aren’t you?’

Table 5.25: Diachrony of selected German and English interrogatives (Hackstein 2004: 175; Seebold 2002; Mallory &amp; Adams 2006: 419f.; and Kroonen 2013: 261, 264)

PIE	PG	Old Germanic	Modern Germanic
*k <sup>w</sup> os <sub>M</sub>	*h <sup>w</sup> az	Got <i>hvas</i>	-
*k <sup>w</sup> is <sub>M</sub>	*h <sup>w</sup> iz	OHG (h)wer	NHG wer
*k <sup>w</sup> eh <sub>2</sub> <sub>F</sub>	*h <sup>w</sup> ō	OE <i>hwā</i> ?OE <i>hū</i>	NE who NE how
*k <sup>w</sup> od <sub>N</sub>	*h <sup>w</sup> at	Got <i>hvat</i> , OHG (h)waz, OE <i>hwæt</i>	NHG was, NE what
*k <sup>w</sup> otero-	*h <sup>w</sup> aþera	Got <i>hvaþar</i> , OHG (h)wedār, OE <i>hwæðer</i>	(NHG <i>weder</i> ‘neither’, NE <i>whether</i> )
*k <sup>w</sup> or	*h <sup>w</sup> ar	Got <i>hvar</i> , OHG (h)war, OE <i>hwær</i>	NHG wo, NE where
*k <sup>w</sup> om	?	Got <i>hvan</i>	-
	?	OHG (h)wanne, (h)wenne, <i>wenno</i> , OE <i>hwanne</i>	NHG wann, (wenn ‘if’), NE when
*k <sup>w</sup> ih <sub>1</sub>	?	OE <i>hwī</i> , <i>hwȳ</i>	NE why
?	?	Got <i>hwaiwa</i> , OHG (h)wio	NHG wie ‘how’

Table 5.26: English, German (own knowledge), Yiddish (Katz 1987: 197; Jacobs et al. 1994: 404, 413-414, *passim*), and Altai Low German interrogatives (Jedig 2014: *passim*); Plautdiitsch forms in square brackets from Nieuweboer (1999)

English	German	Yiddish	Plautdiitsch
who [h-]	wer [-e:ɐ]	ver	veeɐ
how [h-]	wie [-i:]	vi	[vöu]
how much/many	wieviel(e) [-i:-]	vi fil(e)	[vöu fiel]
what	was	vos	vaut
which	welch-er [-ɐ]	velkher	[vöune-]
what kind of	was für [-ɐ] ein- [a-]	vos far a	
where	wo	vu	vuue
wither, where to	wohin	vuhin	vuuehaan
whence, where from	woher [-e:ɐ]		vuueheer, vouheer
when	wann	ven	vənäie
why, how come, for what reason	wieso, weshalb, warum		vuurqm
what for	wozu, für [-ɐ] was	tsu vos, far vos	

- d. **Wie/(Was) du bist schwanger?**  
 how/what you are pregnant  
 ‘You are pregnant?’

English cannot employ the interrogative *how* in these circumstances. The information on Yiddish and Plautdiitsch available to me is insufficient for a comparison.

German *was für ein-* is a complex interrogative similar to English *what kind of*. Interestingly it is still separable as witnessed by the following examples.

- (104) German
- a. **Was\_für\_ein Urlaub ist das?**  
 what.kind holiday is that
- b. **Was ist das für\_ein Urlaub?**  
 what is that kind holiday  
 ‘What kind of holiday is this?’

An analogous situation can be seen in Yiddish.

- (105) Yiddish
- a. **vos\_far\_a yontev iz dos?**  
 what.kind holiday is that
- b. **vos iz dos far\_a yontev?**  
 what is that kind holiday  
 ‘What kind of holiday is this?’ (Jacobs et al. 1994: 413)

For Altai Low German no cognate is attested (but cf. Dutch *wat voor een*). The conjugation of *was für ein-* in German is highly complex and depends on number, gender, and case (Table 5.27). In the plural the interrogative *welch-* ‘which’ substitutes for *ein-* (cf. *eins* ‘one’). Compare the full paradigm of the interrogative *welch-* ‘which’ (Table 5.28). The genitive forms are rare, but are listed for the sake of completeness.

Table 5.27: Conjugation of *was für ein-* ‘what kind of’

Case	SG.M	SG.F	SG.N	PL
NOM	ein-e	ein	ein(s)	Ø/welch-e
ACC	ein-e	ein-en	ein	Ø/welch-e
DAT	ein-er	ein-em	ein-em	Ø/welch-en
GEN	ein-es	ein-er	ein-es	Ø/welch-er

Both *was für ein-/welch-* as well as *welch-* may be used either pronominally or attributively. If used attributively and in the plural, *was für* may be used on its own. In the singular there is the purely pronominal form *was für eins* for the neuter instead of the

Table 5.28: Conjugation of *welch-* ‘which (one)’

Case	SG.M	SG.F	SG.N	PL
NOM	welch-e	welch-er	welch-es	welch-e
ACC	welch-e	welch-en	welch-es	welch-e
DAT	welch-er	welch-em	welch-em	welch-en
GEN	welch-es	welch-er	welch-es	welch-er

attributive form *was für ein*. German *wie viel-*, Yiddish *vi fil(e)*, and Plautdiitsch *vöu fiel* are based on the same underlying pattern as English *how many*. The conjugation of *wie viel-* exhibits the same case markers as the plural forms of *welch-*. While English employs *how much* instead of *how many* for mass nouns, German *wie viel* simply lacks inflection.

In German, Plautdiitsch, Yiddish, and English the personal interrogative shows a small paradigm. The interrogatives meaning ‘what’ do not show case marking.

Table 5.29: German, Yiddish, Plautdiitsch, and English conjugation of the personal interrogative

Case	German	Plautdiitsch	Yiddish	English
NOM	wer [-e:v]	veev	ver	who [h-]
ACC	wen [-e:-]	?	vemen	whom
DAT	wem [-e:-]	veem	vemen	whom
GEN	wessen	?	vemens	whose

Of these four languages only German and perhaps Plautdiitsch preserve four distinct forms, although German *wessen*, which as has an archaic variant *wes*, is increasingly replaced with *von wem* ‘of whom’. German has a parallel paradigm and asymmetry of the definite article or demonstrative *der* ‘that one, the.M.SG’: *der, den, dem, des(sen)*, but *das* ‘that’.

Plautdiitsch *vuurōm* is comparable to German *warum* ‘why’, which is based on MHG *wār + umbe* ‘where + around’. Several more forms in Plautdiitsch such as *vou-bii* (German *wo-bei*) have a locative basis. Unfortunately, only a few forms from Plautdiitsch are attested, which is why German forms are given instead (Table 5.30). English shares some of these formations, e.g. *whereby, thereby, hereby* etc. In one group the *r* is preserved but reanalyzed as belonging to the second element (*wor-um* > *wo-rum*, while in another group the *r* was lost or at least is not present. This seems also to hold for Plautdiitsch, e.g. *vour.op* (German *wor.auf*) and *vou-fōn* (German *wo-von*). Within the first group the vowel following the *r* helped preserve it. Depending on the verb, some of these forms derived from ‘where’ may also just mean ‘what’, which is highly unusual from a typological perspective (Cysouw 2007). Compare, for instance, English *to consist* [*of what*] and German [*wor.aus*] *bestehen*. The development of the meaning of the individual forms is

5 Survey of the grammars of questions in Northeast Asia

Table 5.30: German interrogative and demonstrative paradigms

where	there		here		Explanation
wo(r)	da(r)	hin	hier	her	plain
wo-bei	da-bei	-	hier-bei	-	'at, by, with'
wo-mit	da-mit	-	hier-mit	-	'with'
wo-nach	da-nach	-	hier-nach	-	'to, after'
wo-von	da-von	-	hier-von	-	'from, of'
wo-zwischen	da-zwischen-	-	?hier-zwischen-	-	'between'
wo-hin	da-hin	-	hier-hin	-	'there'
wo-her	da-her	-	hier-her	-	'here'
wo-vor	da-vor	-	?hier-vor	-	'in front'
wo-durch	da-durch	-	hier-durch	-	'through'
wo-zu	da-zu	-	hier-zu	-	'to'
wor.in	d(a)r.in	-	?hier.in	-	'in'
wor.auf	d(a)r.auf	(hi)n.auf	hier.auf	(he)r.auf	'on, up'
wor.unter	d(a)r.unter	(hi)n.unter	hier.unter	(he)r.unter	'under'
wor.über	d(a)r.über	(hi)n.über	hier.über	(he)r.über	'over'
wor.aus	d(a)r.aus	(hi)n.aus	hier.aus	(he)r.aus	'out'
wor.ein	d(a)r.ein	(hi)n.ein	hier.ein	(he)r.ein	'in(to)'
wor.um, war.um	d(a)r.um	hin.um	hier.um	(he)r.um	'about, in order to'

highly idiosyncratic. For example, *wor.über* may either mean 'over what place' but also 'about what'. The close relationship between *da* and *wo* (English *there* and *where*) may be directly traced to Proto-Indo-European where we find the two forms *\*tó-r* and *\*k<sup>w</sup>ó-r* (Mallory & Adams 2006: 419). German *hier* [-i-] (English *here*, Plautdiitsch *hie*, Dutch *hier*) must be a Germanic innovation ultimately based on *\*h<sub>1</sub>ei-* 'this (one)' (Mallory & Adams 2006: 417f.), but it is somewhat obscure (e.g., Seebold 2002).

There are also some parallel forms based on *her* 'here (movement)' (a variant of *hier*) and *hin* 'there (movement), towards' that are also used as preverbs, e.g. *her-kommen* 'to come here', *hin.zu-fügen* 'to add' etc. In German the reanalysis resulted in a few problems such as the fact that there are no separate forms *\*he-* (hence *her.um* > *he.rum* > *rum*), *\*hi-* (hence *hin.ein* > *hi.nein* > *nein*) etc.

### 5.5.3.3 Interrogatives in Slavic

Table 5.31 shows the development of selected Slavic interrogatives over the course of time.

Russian *kotóryj* is a direct cognate of English *whether*, and *či* of *why*. According to Derksen (2008: 172, 227), the second part of Russian *kogdá* is a dative form of PS *\*gódǝ* 'right time' and goes back to PIE *\*g<sup>h</sup>od<sup>h</sup>-o-* (English *good* goes back to PIE *\*g<sup>h</sup>ōd<sup>h</sup>-o-*), based on the stem PIE *\*g<sup>h</sup>ed<sup>h</sup>-* 'join, fit together' (Mallory & Adams 2006: 381). According to Derksen (2008), the final *-li* in PS *\*koli* 'how much' is the Slavic question marker, but

Table 5.31: Diachrony of Slavic interrogatives with selected cognates according to Derksen (2008); PS = Proto-Slavic, OCS = Old Church Slavonic

PIE	PS	OCS	Russian
*k <sup>w</sup> ō-ko-	*kakъ ‘what (kind of)’	kakъ	kakój
*k <sup>w</sup> o-ter-o-	*koteryъ, *kotorъ	kotorъi	kotóryj
*k <sup>w</sup> o-	*kъjъ ‘who, what, which’	kъi	koj
*k <sup>w</sup> o- + *tod	*kъto ‘who’	kъto	kto
*k <sup>w</sup> o- + *g <sup>h</sup> od <sup>h</sup> -o-	*kogъda, *kogъdo ‘when’	kogda	kogdá
*k <sup>w</sup> oli	*koli ‘how much’	koli	(kóli ‘if’)
	*koliko ‘how much’	koliko	(Cz arch. koliko)
*k <sup>w</sup> i/e- + *tod	*čъto ‘what’	čъto	čto
*k <sup>w</sup> iH	*či ‘CONJ’	či ‘because’	(dial. či ‘if, or’)
?	*akъ ‘such as’	jakъ	(Cz jaký ‘which’)

more likely the form simply goes back to PIE \*k<sup>w</sup>oli and is perhaps related to PIE \*k<sup>w</sup>eh<sub>a</sub>li, whence Latin *quālis* (Mallory & Adams 2006: 420). It would be unexpected for a further derivational element to attach to the question marker as in PS \*koliko.

Russian as spoken in Inner Mongolia does not exhibit major differences with respect to Standard Russian. Some dialectal forms from Siberia can be found in Table 5.36 below. In some instances such as *kto* vs. *xto* ‘who’ or *kak* vs. *jak* ‘how’ only phonological differences separate Russian and Ukrainian. In other cases there are different derivations such as in *po-čemu* (DAT) vs. *pó-ščo* (NOM) ‘why’. Only in a few instances are there altogether different interrogatives such as *kogdá* vs. *kolí* ‘when’. The Russian and Ukrainian forms meaning ‘why’ are case forms (DAT, INSTR, GEN, NOM) used with or without prepositions. A preposition can also be found in Russian *ot-kúda*/откуда ‘whence’ (< OCS *ot-* < PIE *h<sub>a</sub>et* ‘away, beyond’, Mallory & Adams 2006: 289ff.). Table 5.33 shows the paradigms of the interrogative pronouns meaning ‘who’ and ‘what’ in Proto-Slavic, Russian, and Ukrainian. Russian *čto*/что ‘what’ has the colloquial pronunciations [tɕ] and an informal variant *čo*/чо [tʃɔ] (e.g., Bai Ping 2011).

The difference between the genitive on the one hand and the nominative on the other, when filling in for the accusative, is connected with the distinction between animate and inanimate meaning (Cubberley 2002: 127). A difference between nominative (e.g. Ukrainian *x-to* ‘who, š-čo ‘what’) and oblique stems (e.g., Ukrainian *k-*, *č-*) is also known from Iranian (see below). Similar to German, the selective interrogative shows extensive paradigms. Russian and Ukrainian also have a distinction between masculine, feminine, and neuter gender but preserve more cases. For reasons of space only some Ukrainian interrogative paradigms will be given in the following (Tables 5.34, 5.35).

As would be expected, most interrogatives in the two pidgin languages are derived from Russian. Table 5.36 shows those interrogatives attested for Taimyr Pidgin Russian. An interesting fact is the frequent use of the oblique forms *kogo* and *čego*. There are three newly formed complex interrogatives. Mednyj Aleut also has some Russian interrogatives (§5.4.3). Apparently, one Nganasan form was borrowed as well.

5 Survey of the grammars of questions in Northeast Asia

Table 5.32: Selected interrogatives from Russian (Wade 2011: passim), Russian as spoken in Inner Mongolia (Bai Ping 2011: passim), and Ukrainian (Pugh & Press 1999: passim)

Meaning	Russian	Russian (China)	Ukrainian
who	kto/кто	ktɔ/кто	xto/хто
whose	čej/чей	tʃej/чей	čuj/чий
what	čto/что	ʃtɔ/что ~ tʃɔ/чо	ščo/що, vŷščo/вищо
who, which	kotóryj/который		kotrŷj/котрий
whither	kudá/куда	ku'da/куда	kudŷ/куди
what (kind of), which	kačój/какой	ka'kɔj/какой	jačŷj/який
how, in what manner	kač/как	kač/как	jač/як
when	kogdá/когда	kač'da/когда	kolí/коли
where	gde/где	gdje/где	de/де
how much/many	skól'ko/сколько	'skɔlka/сколько	skíl'ku/скільки
whence	ot-kúda/откуда		zvidku/звідки, zvidkíl'/звідкіль, (z)vidkiljá/звідкіля
why	po-čemú/почему, za-čém/зачем	za-'tʃjem/зачем	čomú/чому, čohó/чого, pó-ščo/пощо, navŷ-ščo/навищо

Table 5.33: Proto-Slavic, Russian, and Ukrainian interrogative paradigms (Pugh & Press 1999: 178; Shevelov 1993: 961; Sussex & Cubberley 2006: 269f.; Bai Ping 2011: 74, 78)

	PS		Russian		Ukrainian	
Case	who	what	who	what	who	what
NOM	*k-ъ-to	*č-ъ-to	k-to	č-to [ʃtɔ]	x-to	ščo
GEN	*k-ogo	*č-eso/-ьso	k-ogó [kɔ'vɔ]	č-egó [tʃi'vɔ]	k-ohó	č-ohó
DAT	*k-omu	*č-emu	k-omú	č-emú	k-omú	č-omú
ACC	= NOM	= NOM	= GEN	= NOM	= GEN	= NOM
INSTR	*k-ěmь	*č-imь	k-em	č-em	k-ym	č-ym
LOC	*k-omь	*č-emь	k-om	č-ëm	k-ómu, k-im	č-ómu, č-im

Table 5.34: Conjugation of Ukrainian *kotoryj* ‘which’ (Pugh & Press 1999: 180)

Case	M	N	F	PL
NOM	kotrýj	kotré	kotrá	kotrí
GEN	kotróho	id.	kotróji	kotrýx
DAT	kotrómu	id.	kotríj	kotrým
ACC	= NOM/GEN	= NOM	kotróu	= NOM/GEN
INSTR	kotrým	id.	kotróju	kotrýmy
LOC	= DAT, kotrím	id.	= DAT	= GEN

Table 5.35: Conjugation of Ukrainian *jakýj* ‘what kind of’ (Pugh & Press 1999: 180)

Case	M	N	F	PL
NOM	jakýj	jaké	jaká	jakí
GEN	jakóho	id.	jakóji	jakýx
DAT	jakómu	id.	jakíj	jakým
ACC	= NOM/GEN	= NOM	jakú	= NOM/GEN
INSTR	jakým	id.	jakóju	jakýmy
LOC	= DAT, jakím	id.	= DAT	= GEN

Table 5.36: Taimyr Pidgin interrogatives (Stern 2005; 2012: 435ff., 498); some variants were excluded

Form	Meaning	Analysis
syly/сулу	who	Nganasan sili ‘who’
kto/кто	who	= Russian
kakoj/какой	which	= Russian
gde/где	where	= Russian
kuda/куда	whither	= Russian
začem/зачем	why	= Russian
kak/как	how	= Russian
kogda/когда	when	= Russian
kogo/кого	who	Russian GEN/ACC
čego/чего	what	Russian GEN
kudy-mera/куды мера	where	Russian mera/мера ‘measure’
kudy-mesto/куды место	where	Russian mesto/место ‘place’
kakoj storona/какой сторона	whither	Russian storona/сторона ‘side’
akto/акто	who	= dialectal Russian
počto/почто	why	= dialectal Russian
čo/чо	what, why	= dialectal Russian

## 5 Survey of the grammars of questions in Northeast Asia

The phenomenon of one form meaning ‘what’ and ‘why’ is also known from the Iranian languages Sogdian ((ə)ču) and Khotanese (*cu*), see §5.5.3.4.

Only a short list of interrogatives in Chinese Pidgin Russian can be assembled from the material provided by Shapiro (2010). Most forms are of Russian origin, but at least one is from Chinese. The interrogative *mnogo-malo* entirely consists of Russian material but follows the Chinese structure and meaning (Table 5.37).

Table 5.37: Chinese Pidgin Russian interrogatives mentioned by Shapiro (2010); the form in < > was given in Chinese Pinyin

Form	Meaning	Analysis
kaka	why	= Russian kak/как
kakoj	which	= Russian kakoj/какой
<gedao'erli>	which	= Russian kotóryj/который
pocheto	why	= dialectal Russian
mnogo-malo	how much	Russian mnógo/много ‘much’, málo/мало ‘little’, cf. Chinese duōshǎo 多少
šýma	what	= Chinese shénme 什么

### 5.5.3.4 Interrogatives in Iranian

Most interrogatives in those Iranian languages included here are synchronically opaque and their etymologies too complex to be given here in their entirety. But consider the interrogatives in Table 5.38. As can be seen, Sogdian interrogatives have some direct correspondences in, or at least similarities to, Yaghnobi, the only closely related modern

Table 5.38: Sogdian interrogatives (Yoshida 2009: passim) in comparison with Yaghnobi (Geiger 1901: passim; Bielmeier 1989: 482, 484)

Meaning	Sogdian	Yaghnobi	
		Geiger	Bielmeier
when	kaða	kad	kad
which	kataam	kaam	kom, OBL komi kuum, OBL kuumi
where	(ə)kuu	kuu	
who	(ə)ke, OBL (ə)kya	kax, OBL kái	kax, OBL kay
what	(ə)č <u>u</u>	čaa	čo, OBL čoy
why	(ə)č <u>u</u>	čuu	
how many/much	čaaf(ar)	čaaf	čof

language spoken in Tajikistan. Most up-to-date material on Yaghnobi was published in Tajik and has thus to be excluded.

Some of these forms can directly be traced back to Indo-European. For instance, Yaghnobi *kad* ‘when’ goes back to PIE *\*k<sup>w</sup>odéh<sub>2</sub>* and *kuu* ‘where’ to PIE *\*k<sup>w</sup>u(ú)* (or *\*kú-*). The occasional initial vowel in Sogdian is perhaps prothetic (Yoshida 2009: 286). Khotanese has several comparable forms such as *kaama-* ‘which’, *ku* ‘where’, *kye ~ ce* ‘who’, *cu* ‘what, why’, and some additional forms such as *craama-* ‘what kind of’, *ciiyä* ‘when’, *caalsto* ‘whither’, or *canda, cändäka, cerä* ‘how much/many’ (Emmerick 2009: 387, 389, passim). Table 5.39 lists interrogatives from Sarikoli. In order to put them into a proper context, interrogatives from the closely related language Wakhi are listed as well. Remember that these two languages are collectively called Tajik in China but that Tajik is really a variety of Persian.

Table 5.39: Selection of Sarikoli, Wakhi (Gao Erqiang 1985: passim), Tajik, and Persian interrogatives (Windfuhr & Perry 2009: 438); Sarikoli form in square brackets from Xiren Kuerban & Alimujiang Xiren (2015: 88, 120, 162f.); Wakhi forms in square brackets from Bashir (2009: 831); not all variants are listed

Meaning	Sarikoli	Wakhi	Tajik	Persian <sup>11</sup>
which	tʃidum	[kum(d)]	kadom	kodaam
where	ko, ku-dʒui	kum-dʒai	ku, ku-jo	ku, ko-jaa
when	[tʃum]		key	kay
who	tʃoi, OBL -tʃi-	kui	kii	ki
what	tseiz, [tseiz]	tsiz, [čiz]	čii ~ ča-	če(-)
why	[tsarang ‘how’]	[čir]	čaro	čeraa
how many/much	tsund, [tsond, tʃand]	[tsum]	čand	čand

According to Xiren Kuerban & Alimujiang Xiren (2015: 163), Sarikoli *tsond* ‘how many/much’ stands opposed to *tʃand* ‘how many’, which may have been adopted from Tajik. The locative interrogatives contain words meaning ‘place’. All Iranian languages included here preserve the split between the two resonances *k~* and *č~* (or *c~*), the latter of which goes back to *\*k-* as well. Sarikoli, apart from this distinction, has an innovative third type *tʃ~ < k~* that cannot be found in Wakhi or Persian. In Gao Erqiang (1985) it thus shows synchronic variation between *k-*, *ts-*, and *tʃ-*. Khotanese similarly had free variation between *kye* and the more innovative *ce* ‘who’ (Emmerick 2009: 387). Sarikoli *tʃoj* ‘who’ has a special oblique form *tci* that is the basis for case marking (cf. Sogdian and Yaghnobi in Table 5.38).

(106) Sarikoli

- a. *tʃoj*      *a=ta*      *ǰud?*  
 who.NOM ACC-2SG.ACC hit.PST  
 ‘Who hit you?’

<sup>11</sup>Several Persian interrogatives have been borrowed by Moghol (§5.8.3).

5 Survey of the grammars of questions in Northeast Asia

- b. *təw=at*                      *a=tçi*                      *δud?*  
 2SG.NOM=2SG.PST ACC-who.OBL hit.PST  
 ‘Whom did you hit?’
- c. *təw=at*                      *tçi=ri*                      *δud?*  
 2SG.NOM=2SG.PST who.OBL-DAT hit.PST  
 ‘Whom did you give (it) to?’ (Kim 2014: 9)

Gao Erqiang (1985: 35) has the full paradigm as follows: NOM *tfoi*, GEN *tfi-an*, DAT *tfi-ri*, ACC *a-tfi*.

5.5.3.5 Interrogatives in Tocharian

Despite the fact that the two Tocharian varieties are thought to be relatively closely related, there are nevertheless many differences within the system of interrogatives (Table 5.40).

Table 5.40: Selection of Tocharian interrogatives according to Adams (2013) with additional Tocharian A data by Sieg & Siegling (1931: 176–191) and Carling (2009)

	TA	TB
who, what	kus	k <sub>u</sub> se
why	k <sub>u</sub> yal	kā
where	tā	k <sub>u</sub> ta-mem
how, when, if	kupre	kwri, kr <sub>u</sub> i
how much/long	kos	kos, kot
who, which, what kind of	äntsam, antsam	intsu
where	(äntan(n)ene CONJ)	inte, ente
when	(äntāne CONJ)	inte, ente
who, what, which		mäksu
how	mä(n)t	mäkte

This might be additional evidence for Peyrot’s (2010: 144) assumption that a “Proto-Tocharian may have differed more from its daughter languages than is often suggested by superficial similarities between them”, which could be the result of later convergence. The best etymologies for Tocharian interrogatives have been given by Adams (2013), but these are too complex and somewhat too uncertain to be given here in full length.

There is a resonance in *k*(<sub>u</sub>)~ that, as seen before, is a reflex of PIE \**k*<sup>w</sup>~. There are also interrogatives starting with *m*- such as TA *mänt* that might be based on an interrogative stem PIE \**me/o*- (Mallory & Adams 2006: 421). However, Adams (2013) assumes that TB *mäksu* ‘who, what, which’ and *mäkte* ‘how’ as a middle part contain the actual PIE interrogative stem \**k*<sup>w</sup>*i/u*-, preceded by the PIE particle \**men* and followed by different demonstratives or relatives. TB *mäksu* ‘who, what, which’ exhibits a more or

less full paradigm based on person, number, and gender, e.g. *mäksu* ‘NOM.SG.M’, *mäksā<sub>u</sub>* ‘NOM.SG.F’, *mäktu* ‘NOM.SG.N’ (Krause & Thomas 1960: 166). There is some agreement that TA *kus* and TB *k<sub>u</sub>se* < Proto-Tocharian \**k<sup>w</sup>əsē* ‘who’ are similarly combinations of an interrogative with a demonstrative, perhaps PIE \**k<sup>w</sup>i-* + \**so* (e.g., Kim 2012: 38). This is reminiscent of Slavic \**křto* ‘who’ < PIE \**k<sup>w</sup>o-* + \**tod* (see Table 5.33 above). Indo-European had a distinction between three demonstratives, \**so* ‘that one, he’, \**seh<sub>a</sub>* ‘that one, she’, and \**tod* ‘that one, it’ (Mallory & Adams 2006: 417). The difference lies in the fact that Tocharian \**k<sup>w</sup>əsē* contains the first of these, and Slavic \**křto* the last. TB *k<sub>u</sub>se* ‘who’ (TA *kus*) and *k<sub>u</sub>ce* ‘whom’ (TA *kuc*) later had the abbreviated forms *se* and *ce*, respectively (Kim 2012: 38), which is reminiscent of TA *tā* ‘where’ as opposed to TB *k<sub>u</sub>ta-*. In Tocharian B the meaning of both *k<sub>u</sub>se* and *mäksu* encompasses both ‘who’ and ‘what’, which, apart from Baltic languages and Kusunda, is quite exceptional in Eurasia.

## 5.6 Japonic

### 5.6.1 Classification of Japonic

As is by now well established, Japanese is not an isolated language, as was, for instance, claimed by Shibatani (1990: 89). Instead, Japanese is merely the major, but by no means the only, representative of a language family called Japanese-Ryūkyūan or simply Japonic (e.g., Tranter 2012b: 3f.). A simplified classification of Japonic languages may tentatively be represented as in Figure 5.3 (based on Pellard 2009: 264; Chien Yuehchen & Sanada Shinji 2010; Shimoji 2010; Hasegawa 2015: 21ff.), excluding most historically attested and possible Para-Japonic languages. Only those Ryūkyūan languages or dialects mentioned during this section are listed.

The primary split in Japonic is between Japanese and Ryūkyūan. Mainland Japanese constitutes a dialect continuum that can roughly be classified into four larger areas called Eastern, Central, Western, and Kyūshū (Hasegawa 2015: 21f.). The Hachijō dialects and the Okinawan dialect influenced by Ryūkyūan form separate groups in themselves. For reasons of space and lack of sufficient information, a focus will be on Modern Standard Japanese in this study.<sup>12</sup> A special case is Yilan Creole spoken on Taiwan and has thus also been listed separately. Even though the lexicon is mostly based on Japanese, Yilan is a creole language that also exhibits certain influences from Austronesian languages, especially Atayal (Chien Yuehchen & Sanada Shinji 2010). Ryūkyūan languages spoken in the Ryūkyū archipelago may be classified into two main branches, Northern and Southern Ryūkyūan, each of which splits into two branches (Shimoji 2010). Southern Ryūkyūan has also been called Sakishima (Bentley 2008a). Yonaguni is treated as a separate branch of Ryūkyūan by Izuyama (2012) and as a separate subbranch of Southern Ryūkyūan by Bentley (2008a: 242), but is often included within the Macro-Yaeyama subbranch of Southern Ryūkyūan, the other branch of which is Miyako. Northern Ryūkyūan can be divided into Amami and Okinawan. There is a large amount of variation among Ryūkyūan. According to Lawrence (2012: 380), there are 35 “dialects” within Miyako and

<sup>12</sup>A *Handbook of Japanese Dialects* has been announced by De Gruyter for 2019.

5 Survey of the grammars of questions in Northeast Asia

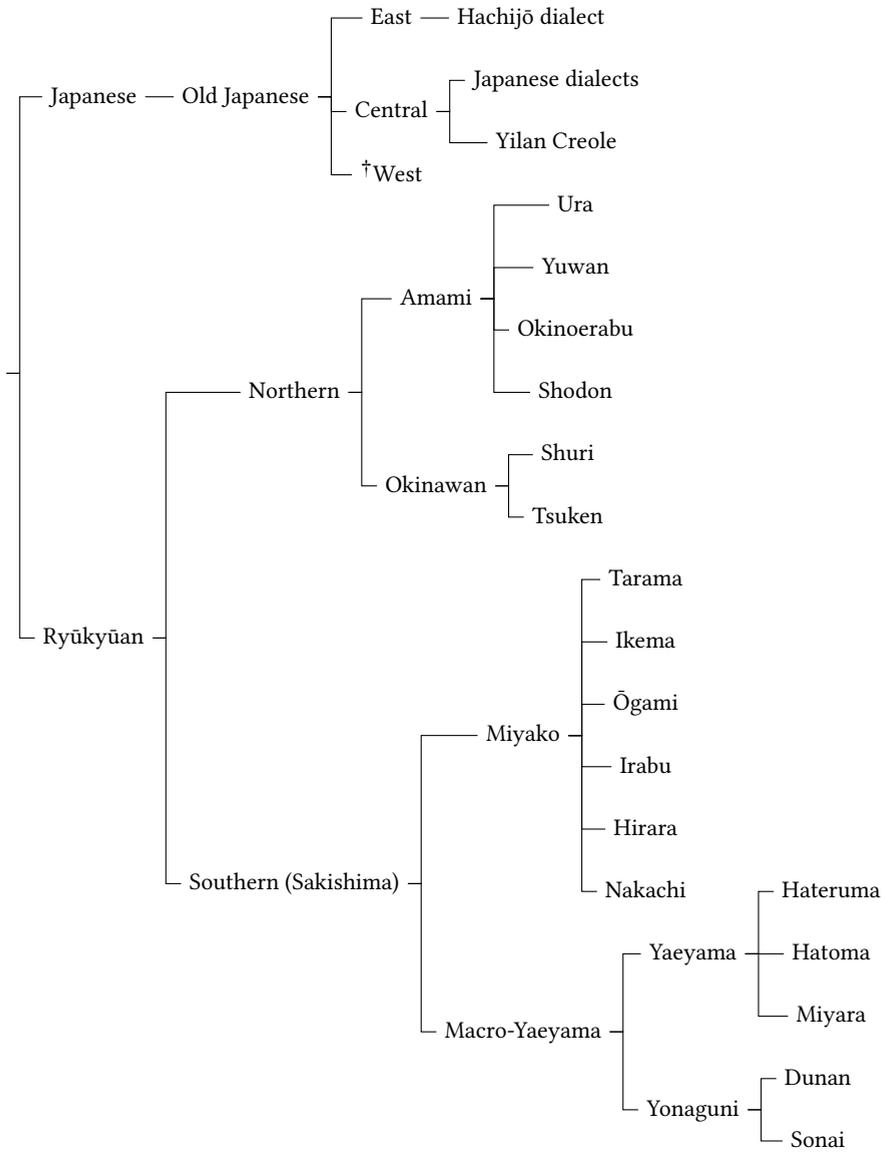


Figure 5.3: Classification of Japonic

20 within Yaeyama alone. Of course, a classification into languages and dialects is difficult and even somewhat spurious. But clearly the Ryūkyūan Islands can be considered a treasure trove of linguistic diversity, of which only some parts can be included in this chapter. As is common practice today in the study of Ryūkyūan, the place name will also indicate the language spoken at that place, i.e. Irabu on Irabu island etc.

### 5.6.2 Question marking in Japonic

Tranter & Kizu (2012: 295) give a good summary of question marking strategies in modern Standard Japanese.

Questions of all types including *wh*-questions and *yes/no* questions are expressed by a change in intonation and the addition of a particle at the end of the sentence: familiar-style *-no*,  $-\emptyset$ , or polite-style *-ka*. Soliloquy-type questions that do not necessarily require a response from a hearer use *-kana(a)* or *-kashira* (female). There is no change in word order, and no fronting in *wh*-word questions. Questions that present alternatives, including those that ask a question in an affirmative form with a negative alternative of the same situation, have the structure of two separate questions.

The speech level differences are not as strongly developed as they are in Koreanic questions (§5.7.2). The default and polite question marker in Japanese is the sentence-final and possibly enclitic particle *ka* か. The marker can be found in polar, alternative, and content questions.

(107) Japanese

- a. *dare.ka ga kimashita ka?*  
 someone NOM come.PST.POL Q  
 ‘Did someone come here?’ (Hasegawa 2015: 104)
- b. *sono hon wa omoshiroi desu ka tsumaranai desu ka?*  
 that book TOP interesting COP Q boring COP Q  
 ‘Is that book interesting or boring?’ (Hinds 1984: 159f.)
- c. *o-namae wa nan desu ka?*  
 HON-name TOP what COP Q  
 ‘What is (your) name?’ (own knowledge)

The same marker also appears at the end of what seem to be focus questions. The following two examples were elicited from a native speaker living in Germany in November 2015. The glossing follows Hasegawa (2015).

(108) Japanese

- a. *ashita wa gakkō ni iki-masu ka?*  
 tomorrow TOP school to go-NPST.POL Q  
 ‘Is it tomorrow that you are going to school?’

5 Survey of the grammars of questions in Northeast Asia

- b. *ashita, gakkō ni wa iki-masu ka?*  
 tomorrow school to TOP go-NPST.POL Q  
 ‘Is it to school that you are going to tomorrow?’

Similar to Korean (§5.7.2) and Wutun (§5.9.2.1), it apparently is the topic marker *wa* that follows the focused or perhaps rather topicalized element while the sentence otherwise is identical to a plain polar question. Japanese has a special way of forming topic questions that contain the same topic marker *wa* but have a truncated form.

- (109) Japanese  
*anō, kyōdai wa?*  
 uh siblings TOP  
 ‘So, do you have any brothers or sisters?’ (Hinds 1984: 166)

In Old Japanese, the particle *ka* already existed but differed from the modern Japanese one in its syntactic behavior. According to Vovin (2009: 1220), it was present in both Eastern and Western Old Japanese and has the same scope as in modern Japanese. But, in contrast to the strict sentence-final position today, the particle could appear in other positions as well. Apparently, the particle also marked focus questions and attached to the focused element in both focus and content questions.

- (110) Old Japanese
- a. 囊伽多佐例  
*ta=ka ta-sar-e?*  
 who=Q EMP?-go.away-EV  
 ‘Who goes away?’
- b. 今夜可君之我許來益武  
*KŌ YŌPI=ka KYIMYI-NKA WA-Nkari K-YI-[i]mas-am-u?*  
 this night=Q lord-POSS 1SG-DIR come-INF-HON-TENT-ATTR  
 ‘Is it *tonight* that (my) lord will come to me?’ (Western; Vovin 2009: 1220, 1225)

Typologically, this is a change similar to the one observed from Middle Mongol to modern Mongolian (see §5.8.2).

In Eastern Old Japanese *=ka* is attested as a marker for polar, focus, and content questions and triggers *kakari musubi* ‘focus concord’ (see further below): It “forces the main verb to take an attributive suffix, regardless of whether it follows or precedes the verb” (Kupchik 2011: 834).

- (111) Old Japanese (Eastern)
- a. 安杼加安我世牟  
*aNtō=ka a-Nka se-m-u?*  
 what=Q 1SG-POSS do-TENT-ATTR  
 ‘What should I do?’

- b. 於不世他麻保加  
*opuse-tamap-o=ka?*  
 assign.INF-HON-ATTR= Q  
 ‘Has (the emperor) given (me) the order?’
- c. 夜麻尔可祢牟/毛夜杼里波奈之尔  
*yama-ni=ka ne-m-u mwo / yaNtör-i pa na-si-ni?*  
 mountain-LOC=Q sleep-TENT-ATTR FOC / lodge-N TOP not.exist-FIN-LOC  
 ‘Shall (I) sleep *in the mountains* since there is no lodging (here)?’ (Kupchik 2011: 834, 835)

According to Vovin (2009: 1229), the particle has a cognate in Ryūkyūan languages and can be traced back to Proto-Japonic.

The interrogative particle *ka ~ ga* (< \*-*N ka*) is well attested in both Old Ryūkyūan and modern dialects. However, as far as I can tell, Ryukyuan *ka ~ ga* appears exclusively in *wh*-questions [CQ]. Thus, in all probability, WOJ [Western Old Japanese] *ka* in general questions [PQ] represents a Japanese innovation, and we should reconstruct PJ [Proto-Japonic] \**ka*, interrogative particle in *wh*-questions.

An example from Old Ryūkyūan is the following:

- (112) Old Ryūkyūan  
 けおわのかしよらしよ  
*keo wa no=ka s-i-yor-asiyo?*  
 today TOP what=Q do-INF-exist-SUP  
 ‘What would (they) do today?’ (Vovin 2009: 1229)

Old Korean had a similar marker *-ka* 去 that might be somehow related to the Japonic form (§5.7.2). But as we will see later in some Ryūkyūan languages, there is also the possibility that the marker is the result of a language internal development from a focus marker.

Japanese exhibits an instance of the grammaticalization from nominalization to question marker through ellipsis of the following copula and original question marker.

- (113) Japanese  
*doko e iku no (desu ka)?*  
 where ALL go N>Q COP Q  
 ‘Where are you going?’ (Hinds 1984: 163)

The suffix *no* originally may have been the genitive case marker (Shibatani 1990: 258). See §5.1.2 on Ainuic and below on Ryūkyūan for similar developments from nominalizer to question marker that may be due to contact with Japanese. According to Hasegawa (2015: 297) *no* adds “various nuances, typically softening the locution when addressing an interlocutor. It is, therefore, considered mildly feminine even though male speakers

also use this particle.” The most aberrant Japanese dialect, *Hachijō*, has a marker *kai* that was written attached to a preceding word or with a hyphen and translated with Japanese *か*. Presumably, it is either a particle or an enclitic. Content questions seem to remain unmarked (Kokuritsu Kokugo Kenkyūjo 1950: 130, 208). The Tsuruoka dialect in northern Honshū marks polar questions with *ga* and content questions with *na* (Matsumori & Takuichiro 2012: 323, 325). In the Ei dialect in southern Kyūshū, both polar and content questions take the marker *ka* or its formal variant *kana* (Matsumori & Takuichiro 2012: 342).

**Yilan Creole** has the two optional sentence-final markers *ga* and *no*, corresponding to Japanese *ka* and *no*, respectively. As opposed to Japanese *ka*, *ga* apparently does not appear in content questions, which remain unmarked. This may be due to influence from Atayal or Chinese. Polar questions generally have a rising intonation.

(114) Yilan Creole

- a. *kore ga nani?*  
this TOP what  
'What (is) this?'
- b. *anta kyolai aru ga?*  
2SG sibling have Q  
'(Do) you have siblings?'
- c. *anta no hoyin ga qalux no?*  
2SG GEN dog TOP black Q  
'(Is) your dog black?'<sup>13</sup>
- d. *anta teykan 'suw?*  
2SG chair heavy  
'(Is) your chair heavy?' (Peng Qiu 2015: 52, 54, 55)<sup>14</sup>

Yilan Creole questions thus behave very similarly to those in Japanese but have a slightly different form and semantic scope.

The last question marker mentioned by Tranter & Kizu (2012: 295) as quoted above is *-ka.na(a)* or *-ka.shira*, formerly used in women's speech, employed for questions to oneself. According to Hasegawa (2015: 294) *ka.shira* "expresses uncertainty and curiosity" and has been translated as 'I wonder'. As we will see below, Ryūkyūan languages have similar markers containing an element *-ka- ~ -ga-* that was translated in the same way.

Both Eastern and Western Old Japanese had another question marker *ya* found in polar and focus questions. Its behavior in these two dialect groups is rather similar, but there are minor differences. For Eastern Old Japanese we have the following description by Kupchik (2011: 832):

<sup>13</sup>The words *hoyin* 'dog' and *qalux* 'black' have been borrowed from Atayal.

<sup>14</sup>The words *teykan* 'chair' and *'suw* 'heavy' derive from Atayal.

When in the sentence-final position, it follows the copula *tō*, the defective verb *tō* ‘think,’ or the evidential form of the verb. The examples with the evidential are used to make ironic questions [...]. When this particle is fronted to a pre-verbal position, the verb form must take the attributive suffix [...]. Unlike WOJ, where *ya* is amply attested directly after the final form of a verb or the final exclamative *-umo* (Vovin 2009: 1211), such usages are unattested in EOJ.

In Western Old Japanese, the non-final position – presumably found in focus questions – also accompanies the attributive form of the verb. In case it is sentence-final – in polar questions – it may follow final, evidential, and exclamative forms, but not attributive ones (Vovin 2009: 1211).

## (115) Old Japanese (Eastern)

- a. 宇恵古奈宜/賀久古非牟等夜  
*uwe kwo-na-N-kyi / Nka-ku kwopiy-m-u tō=ya?*  
 sow.INF dim-water-LOC-leeks / be.thus-INF long.for-tent-FIN think/say=Q  
 ‘Do (you) think (I) love the sowed water leeks so much?’
- b. 阿須由利也/加曳我伊牟多祢牟  
*asu-yuri=ya / kaye-Nka muta ne-m-u?*  
 tomorrow-ABL=Q / reed-POSS together.with sleep-TENT-ATTR  
 ‘From tomorrow shall (I) sleep together with the reeds?’ (Kupchik 2011: 832)

## (116) Old Japanese (Western)

- a. 儺波企箇輸擲  
*na pa kyik-as-u=ya?*  
 2SG TOP ask-HON-FIN=Q  
 ‘Shall (I) ask you?’
- b. 枳彌波夜那祇  
*kyimyi pa=ya na-kyi?*  
 lord TOP=Q NEG-ATTR  
 ‘Don’t (you) have a lord?’ (Vovin 2009: 1211, 1215)

Similarly to the particle *ka*, Vovin (2009: 1219) assumes that *ya* has cognates in Ryūkyūan and that it can be traced back to Proto-Japonic.

The cognates *ya ~ yaa* of the Western Old Japanese interrogative particle *ya* are well attested in modern Ryukyuan dialects, although in most dialects *ya ~ yaa* have the function of a confirmation seeker, like MdJ *ne*, and not an interrogative particle. As far as I can tell, *ya ~ yaa* occurs only in sentence-final position.

But according to Shinzato (2015: 305), the Old Japanese marker rather corresponds to the Ryūkyūan question marker (*y*)*i*, on which see below. Whether Ainu *ya* may be compared remains an open question, but it may well have been borrowed from older

stages of Japanese (§5.1.2). A sentence-final particle *ya* in Standard Japanese is usually accompanied by falling intonation and does not express questions (Hasegawa 2015: 298).

In Standard Japanese there is another sentence-final particle *kke*, the function of which Hayashi (2010a: 2687) explains as follows:

Thus, unlike *ka* and *no*, *kke* makes implicit reference to knowledge or information previously held by the speaker and shared with the addressee, but which the speaker has somehow forgotten or is unsure about. The particle then serves to enlist collaborative participation of the addressee in the process of regaining that knowledge/information.

- (117) Japanese  
*are ichi-nen deshita kke:?*  
 EXCL one-year COP.PST.POL Q

‘Wait, is (your visa valid) for one year?’ (Hayashi 2010a: 2687, simplified)

There is also a special marker *tte* for echo questions, which is a variant of the quotative marker *to* used in casual speech. But *to* cannot function as a sentence-final particle (Hasegawa 2015: 310f.).

- (118) Japanese  
*dare deshita tte?*  
 who COP.PST.POL QUOT>Q

‘Who did you say it was?’ (Hinds 1984: 165)

According to Hinds (1984: 165), the marker has its origin in an ellipsis of the subsequent speech act verb followed by the question marker *ka*.

In a comparative study of question-response sequences in ten different languages, Japanese had the highest ratio of polar questions (85%), as opposed to content questions (15%). There was only one alternative question. But 39% percent of the polar questions had a declarative form and 30% were actually tag questions (Hayashi 2010a: 2686). There were three different tag question markers, *janai*, *deshō*, and *yo ne*. The first is a negative copula *ja-nai* ‘COP-NEG’ and can roughly be translated as ‘isn’t it?’. It has the shorter version *jan* and a more polite variant *janai desu ka*. The tag marker *deshō* and its less polite variant *darō* are actually so-called conjectural copula forms meaning ‘probably be’ (Hasegawa 2015: 80) and “ask for the addressee’s confirmation to the speaker’s conjecture” (Hayashi 2010a: 2689). They roughly correspond to English tag questions such as ‘is it?’. The last form *yo ne* is a combination of two different markers the function of which goes well beyond the marking of questions (see Hasegawa 2015: 299ff.). According to Hayashi (2010a: 2690), “these particles are used sentence-finally to make an assertion while seeking confirmation/agreement to it from the addressee.” In combination, *yo ne* was translated as ‘don’t you think?’ But *ne* can also mark questions on its own. It has a variant *na* that is usually used by men.

- (119) Japanese  
*ii tenki da na?*  
 good weather COP Q  
 ‘It is a fine day, isn’t it?’ (Hasegawa 2015: 296)

Whether this might be a cognate of a question marker found in several Ryūkyūan languages remains unclear to me.

The marking of questions in **Ryūkyūan languages** is less well described than for Japanese. In general, there are similar patterns with sentence-final particles, but in some languages there are question suffixes and the pattern of question marking may be quite complex. In *Ura* (spoken on Amami Ōshima) polar questions, for instance, there is either a rising intonation or a simple sentence-final clitic =*na* ~ =*nja*.

- (120) *Ura*  
*kuri=ja hon=na?*  
 this=TOP book=Q?  
 ‘Is this a book?’ (Shigeno 2010: 27)

There is an additional marker for “self-questions”, the semantic scope of which was not given. It might belong to other forms meaning ‘I wonder’, e.g. Japanese *-ka.na*.

- (121) *Ura*  
*an ?cju=ja taru=kai?*  
 that person=TOP who=Q  
 ‘Who is that person?’ (Shigeno 2010: 27)

Shigeno (2010) does not further specify whether content questions receive a special marking or not, but among his examples there are the markers =*joo* (in CQ), =*kana* (in CQ), and =*ja(a)* (in PQ), that were glossed as question markers but not further explained.

- (122) *Ura*
- a. *wan=ga kak-ju-Ø-n=ja?*  
 1SG=NOM write-IPFV-(NPST)-ADN=Q  
 ‘Should I write?’
  - b. *an ?cju=nkja=ja icu=raga kuma=nan ur-i=joo?*  
 that person=APPR=TOP when=ABL here=LOC1 exist-NPST=Q  
 ‘What time did those people get here?’
  - c. *nan=cjuu=no=kana?*  
 what=QUOT=GEN=Q  
 ‘How should (I) express this?’ (Shigeno 2010: 20, 23, 30)

The enclitic =*ja(a)* is formally identical to the topic and persuasion markers.

## 5 Survey of the grammars of questions in Northeast Asia

A much better description can be found for the closely related language **Yuwan** (also spoken on Amami Ōshima). In this variety, the marking of questions is much more complicated and displays a typologically very interesting pattern. Similar to Ura, polar questions are either expressed with rising intonation or an enclitic =*na*.

- (123) Yuwan  
*uro=o koow-an=na?*  
2SG=TOP buy-NEG=Q  
'Don't you buy it?' (Niinaga 2015: 337)

But questions in Yuwan may also be expressed by means of affixes. The information is insufficient to decide about the distribution of these three different marking strategies.

- (124) Yuwan  
*uro=o koo-ju-mi?*  
2SG=TOP buy-IPFV=Q  
'Do you buy it?' (Niinaga 2015: 337)

Altogether, there are the three suffixes, *-mi* for polar questions, *-u* for content questions, and *-ui* for focus questions. If that is not enough, the latter two suffixes are not used in isolation but obligatorily combine with focus markers that are specific to the question types, i.e. =*ga* in content and =*du* in focus questions.

The clitic =*du* cannot appear with *-u*, while =*ga* cannot appear with *-ui*. This kind of phenomenon, where the presence of a focus clitic correlates with the type of verbal inflection, is known as *kakari musubi* in Japanese linguistics (Niinaga 2010: 75)

The phenomenon called *kakari musubi* will be discussed in more detail below. As seen above, neutral polar questions take no focus marking.

- (125) Yuwan
- a. *kuri=ba tu-ju-mi?*  
this=ACC take-IPFV-Q  
'Will (you) take this?'
  - b. *nuu=ba=ga tu-jur-u?*  
what=ACC=FOC take-IPFV-Q  
'What will (you) take?'
  - c. *kuri=ba=du tu-jur-ui?*  
this=ACC=FOC take-IPFV-Q  
'Will (you) take *this*? (Niinaga 2010: 76f.)

From a diachronic perspective the content and focus question markers perhaps contain the same element *-u*. The element *-i* possibly has a cognate in Shuri *-i(i)* (or perhaps *=ji*). Clearly, Yuwan *-mi* is cognate with Shuri and Tsuken *-mi*. It has been proposed that these also contain an actual question marker *-i ~ -i*. The focus marker *=du* may also appear in declarative sentences while *=ga* is restricted to content questions (Niinaga 2010: 75). The three question markers exhibit an interesting interaction with polarity and tense (Table 5.41). In the past tense the question markers attach to the “declarative” (past) marker *-tar*, the loss of the *r* before consonants is regular. In non-past tense, on the other hand, the question markers replace the declarative *-i* (perhaps cognate with Shuri *-i* ‘PRS.PTCP’).

Table 5.41: Declarative and question markers in Yuwan (Niinaga 2010: 64)

	Category	Assertion	Negation
NPST	DECL	<i>-i</i>	<i>-an</i>
	PQ	<i>-mi</i>	<i>-ami</i>
	FQ	<i>-ui</i>	-
	CQ	<i>-u</i>	-
PST	DECL	<i>-tar</i>	<i>-an-tar</i>
	PQ	<i>-ta-mi</i>	<i>-an-ta-mi</i>
	FQ	<i>-tar-ui</i>	<i>-an-tar-ui</i>
	CQ	<i>-tar-u</i>	<i>-an-tar-u</i>

In the non-past the polar question marker has a special negative form *-ami* as opposed to the plain negative *-an*. Negative forms of *-ui* and *-u* apparently only exist in the past tense.

Another question marker *=ga(i)* is always used in combination with the suppositional enclitic *=daroo*. The following sentence was translated as a tag question by Niinaga.

- (126) Yuwan  
*an ?cjoo sjensjee=ja ar-an=daroo=ga(i)?*  
 that person.TOP teacher=TOP COP-NEG.NPST=SUPP=Q  
 ‘(I) suppose that that person is not a teacher, is that right?’ (Niinaga 2010: 73)

Niinaga (2010: 72) also used the gloss ‘confirmative question’ for *=ga(i)*. Another enclitic *=jəə* “is used only with intentional inflection to confirm the hearer acknowledges the intention of the speaker” (Niinaga 2010: 72).

- (127) Yuwan  
*waŋ=ga ik-joo=jəə?*  
 1SG=NOM go-INT=Q  
 ‘I will go, right?’ (Niinaga 2015: 329)

5 Survey of the grammars of questions in Northeast Asia

Van der Lubbe & Tokunaga (2015) give an overview of two dialects spoken on **Oki-noerabu** among the Amami islands, Masana in the west and Kunigami in the east. But most examples for questions are from Masana. **Masana** has the same enclitic =*na* ~ =*nja* for polar questions as several languages mentioned above, but in content questions the same form =*joo* as in Ura is found.

(128) Okinoerabu (Masana)

a. ?*tia-η=nja*?

know-IND=Q

‘Do you know?’

b. ?*uda=gatfi=joo*?

where=DIR=Q

‘Where are you going?’ (van der Lubbe & Tokunaga 2015: 353)

The dubitative suffix *-ra* usually combines with the focus marker =*ga* and was translated as ‘I wonder if’ but is not a question marker in the strict sense. Another dubitative marker *-ro* on the other hand is “used to ask questions in a less direct way” (van der Lubbe & Tokunaga 2015: 357).

(129) Okinoerabu (Masana)

*kiba-ti mee-ro*?

work.hard-MED exist.HON-DUB

‘Are you working hard?’ (a greeting) (van der Lubbe & Tokunaga 2015: 357)

Exactly the same description was given for three other markers. The enclitic =*kaja* could be related to Shuri =*gajaa*. Both can be found in content questions, e.g. *taru=kaja*? ‘Who would that be?’ (van der Lubbe & Tokunaga 2015: 362). The origin of the other two (PQ =*sa*, CQ =*do*) remains unclear for now. According to van der Lubbe & Tokunaga (2015: 361) “in the past tense, the medial converb is used rather than the past tense suffix *-ta-*.” In **Kunigami** and other varieties in the eastern part, the verbal suffix *-jee* is employed instead of the enclitic =*na* ~ =*nja*. This might be a cognate of Yuwan =*jəə* and Ōgami *-εε* that we will soon encounter, e.g. *kuruma ?a-jee*? ‘car COP-Q’ ‘Is there a car?’ (van der Lubbe & Tokunaga 2015: 362).

**Shuri** (or Okinawan) as spoken on Okinawa has several question markers and displays strong similarities to other languages mentioned thus far. There is a particle *naa* that has a short vowel in Ura, Yuwan, Tsuken, Tarama, and Ikema and in these languages has sometimes been analyzed as enclitic, sometimes as freestanding particle. It has been translated as a tag question by Miyara (2015), but may also be a plain polar question marker.

(130) Shuri

*kamadee=ga mango tfuku-ta-n=naa*?

PN=NOM mango grow-PST-IND=Q

‘Kmadee grew mangoes, didn’t he?’ (Miyara 2015: 394)

But Shuri also has an interrogative verb morphology. In some cases it is not entirely certain that we are not dealing with enclitics instead, but for purposes of comparison all forms have been given as suffixes. Similar to Yuwan, there is a polar question suffix *-mi*, but content questions take the suffix *-ga*. According to Uemura (2003: 95), as well as (Arakaki 2003: 181f.), however, the actual question marker for polar questions is *-i* and the *-m* originally was an affirmative, declarative, or indicative marker that has the form *-N* in other contexts. According to Arakaki (2010), the suffix *-N* is an evidential marker for “direct evidence”. As opposed to Yuwan, which uses the plain negative *-an* and the interrogative negative *-ami*, Shuri retains its original form in the negative, i.e. *-(r)an-i*. While in Yuwan the new polar question marker simply attaches to the past tense form (*-ta-mi*), Shuri has an amalgamated form *-ti(i)* that in all likelihood goes back to a combination of the past tense marker *-ta* and the interrogative *-i*. However, Uemura (2003: 145) and Arakaki (2015: 67) seem to suggest a combination of the past participle and the question marker instead. In content questions, *-ga* takes the last position, is fully analyzable, and always replaces the indicative ending *-N*. Table 5.42 gives an overview of Shuri verb forms with a focus on interrogative verb morphology.

Table 5.42: Shuri verb forms illustrated with the verbs ‘*uki-* ‘to wake up’ and ‘*kac-* ‘to write’ according to Arakaki (2003: 180f., *passim*); partly reanalyzed (cf. Uemura 2003)

	Vowel stem	Consonant stem
PRS.PTCP	‘ <i>uki-i</i>	<i>kac-i</i>
NPST-IND	‘ <i>uki-ju-N</i>	<i>kac-u-N</i>
NEG (NPST)	‘ <i>uki-ran</i>	<i>kac-an</i>
NPST.NEG-Q (PQ)	‘ <i>uki-ran-i</i>	<i>kac-an-i</i>
NPST-IND.Q (PQ)	‘ <i>uki-ju-mi</i>	<i>kac-u-mi</i>
NPST-Q (CQ)	‘ <i>uki-ju-ga</i>	<i>kac-u-ga</i>
NPST.NEG-Q (CQ)	‘ <i>uki-ran-ga</i>	<i>kac-an-ga</i>
PST.PTCP	‘ <i>uki-ti</i>	<i>kac-i</i>
PST-IND	‘ <i>uki-ta-N</i>	<i>kac-a-N</i>
NEG-PST-IND	‘ <i>uki-ran-ta-N</i>	<i>kac-an-ta-N</i>
NEG-PST.Q (PQ)	‘ <i>uki-ran-ti</i>	<i>kac-an-ti</i>
PST.Q (PQ)	‘ <i>uki-ti(i)</i>	<i>kac-i(i)</i>
PST-Q (CQ)	‘ <i>uki-ta-ga</i>	<i>kac-a-ga</i>
NEG-PST-Q (PQ)	‘ <i>uki-ran-ta-ga</i>	<i>kac-an-ta-ga</i>

Uemura (2003: 95) furthermore mentions the partly suppletive copula forms ‘*ja-N* (affirmative), ‘*ja-mi* (polar question), and ‘*ra-ran-i* (negative polar question). Consider some examples with interrogative verb morphology.

5 Survey of the grammars of questions in Northeast Asia

(131) Shuri

a. *kamadee=ga mangoo tfuku-ju-mi?*

PN=NOM mango grow-PRS-Q

‘Will Kamadee grow mangoes?’

b. *taa=ga mangoo tfuku-ta-ga?*

who=NOM mango grow-PST-Q

‘Who grew mangoes?’ (Miyara 2015: 393, 394)

Whether the suffix *-i(i)* seen above has to be differentiated from the particle *=ji* found in focus questions, remains unclear.

(132) Shuri

*kamadee=ga=du mangoo tfuku-ju-ru=ji?*

PN=NOM=FOC mango grow-PRS-NIND=Q

‘Is it Kamadee who grows mangoes?’ (Miyara 2015: 394)

According to the description by Arakaki (2003: 181f.), the question marker *-i(i)* attaches directly to the verb stem and replaces the usual past tense ending *-a-N*.

(133) Shuri

a. *waN=nee tigami kac-a-N?*

1SG=TOP letter write-PST-IND

‘I wrote a letter.’

b. *‘jaa=ja tigami kac-ii?*

2SG=TOP letter write-Q

‘Did you write a letter?’ (Arakaki 2003: 181)

However, if *-ti(i)* indeed stems from *-ta + -i* (or *-ti + -i*), perhaps *-i(i)* can be analyzed as *-a + -i* (or *-i + -i*). The occasional long vowel (*-tii, -ii*) in Arakaki’s (2003) and Uemura’s (2003) data might be a reflex of this. In (132) above, the focus marker *=du* requires the non-indicative ending *-ru* on the verb. The Yuwan verbal ending *-ui*—combined with *=du* as well—possibly contains a cognate of Shuri *-ji* (or perhaps *-i(i)*). Content questions in Yuwan only have the ending *-u*. In Shuri, if the focus marker *=ga* is present, again identical to the question marker in content questions, the verb takes the question or dubitative marker *-ra*. This pattern can be found in both content and focus questions. See below on *kakari musubi* for further information on this phenomenon.

(134) Shuri

a. *nuu tfi-yu-ga?*

what wear-PRS-Q

‘What do you wear?’

- b. *nuu=ga tʃi-yu-ra?*  
 what=FOC wear-PRS-Q  
 ‘What do you wear?’ (Nagano-Madsen 2015: 204)
- c. *kamadee=ga=ga mangoo tʃuku-ju-ra jaa?*  
 PN=NOM=FOC mango grow-PRS-Q Q  
 ‘Is it Kamadee who will grow mangoes?’ (Miyara 2015: 394)

Apart from all the different forms mentioned, the last example has yet another particle *jaa*, originally glossed as ‘I wonder’, that can also appear as a part of the complex form *ga-jaa*. As noted above, it may be related to the form *ya* in Old Japanese. The first element is unlikely to be the content question marker *-ga* because *gajaa* can also appear in polar questions. The description is insufficient to give a good summary here but (*ga*)*jaa* appears in focus and content questions.

- (135) Shuri  
*kamadee=ja nuu tʃuku-ju-gajaa?*  
 PN=NOM what grow-PRS-Q  
 ‘Kamadee is going to grow what?’ (Miyara 2015: 395)

As opposed to other Ryūkyūan languages the marker *-ka* does not mark neutral questions but rather suggestions.

- (136) Shuri  
*?ari=ga ?i-i-fe=e tʃik-an-ka?*  
 3SG=NOM say-PRS-N-TOP? listen-NEG-SGS  
 ‘Shall we not listen to him?’ (Miyara 2015: 395)

Intonation in Shuri is exceptionally well described and too complex to go into every detail here (see Nagano-Madsen 2015). Several important points have been summarized as follows:

In Japanese, both yes-no and wh-questions are accompanied by final rising pitch. In Okinawan, neither yes-no questions nor wh-questions are accompanied by final rising pitch. Like a yes-no question, Okinawan wh-question has intonation composed of its lexical accent type unless the verb is immediately preceded by a wh-word. When a verb is immediately preceded by a wh-word, the lexical accent of the verb is usually strongly compressed or rather deleted. [...]

Although the most usual form of forming interrogatives in Okinawan is with a mood suffix, it is not impossible to make an utterance that has (declarative) indicative mood suffix +N, which is produced with a final rising pitch. Furthermore, it is quite common to form an interrogative with the sentence-final question particle *na*, which is also produced with a final rising pitch. (Nagano-Madsen 2015: 209)

**Tsuken** (spoken on Tsuken island close to Okinawa) has a polar question marker *-mi* that probably is related to the marker *-mi* in Shuri or *-mi* in Yuwan. At first glance, the question marker replaces the declarative ending in the following examples in a non-past tense. But in fact, *-mi* must go back to *\*-n-i* as in Shuri and Yuwan.

- (137) Tsuken
- a. *ʔjaa=ga kak-u-n.*  
2SG=NOM write-NPST-DECL  
'You will write.'
  - b. *ʔjaa=ga kak-u-mi?*  
2SG=NOM write-NPST-Q  
'Will you write?' (Matayoshi 2010b: 92)

But there is also a cognate of the marker *=na ~ =nja* in Yuwan and other languages that enclitically attaches to the sentence. It does not replace the declarative marker but rather attaches to it.

- (138) Tsuken
- kuruma=kara si-sa-n=na?*  
car=ABL come-PST-DECL=Q  
'Did you come by car?' (Matayoshi 2010b: 102)

The distribution between the two markers also remains unclear in Tsuken but probably is connected to the verb ending. Content questions have a marker *=ga* that, as in Shuri, looks suspiciously similar to the focus marker *=ga* (Matayoshi 2010b: 102). A connection with the nominative/genitive *=ga* seems unlikely.

- (139) Tsuken
- taa=ga sa=ga?*  
who=NOM do=Q  
'Who does?' (Matayoshi 2010b: 94)

There is no example in which the plain focus marker *=ru* is found in a question, which does not mean, however, that this is impossible. The same is true for the focus marker *=du* in the language Tarama.

**Tarama** (spoken on Tarama and Minna among the Miyako islands) otherwise has a straightforward pattern with *=na* found in polar questions and *=ga* in content questions. Again, the optional focus marker in content questions is identical in form with the question marker.

- (140) Tarama
- a. *kure=e kam=nu sima=na?*  
this=TOP god=GEN island=Q  
'Is this an island of god?'

- b. *naa=ju=ba nuu=ti=ga īi=ga?*  
 name=ACC=TOP what=QUOT=FOC say=Q  
 ‘What is your name?’ (Aoi 2015: 417)

There are also examples where there is only one marker with the form =*ga*. Aoi glosses the form as question, but it might well be the focus marker.

- (141) Tarama  
*nuu=ga si-tar?*  
 what=?Q do-PST  
 ‘What happened (with you)?’ (Aoi 2015: 419)

**Ikema** (spoken on Ikema, Irabu, and Miyako among the Miyako islands) also has the two question markers =*na* (PQ, FQ) and =*ga* (CQ). But, as opposed to Yuwan, for instance, the focus marker =*du* appears not only in focus but also in content questions.

- (142) Ikema  
 a. *husi=nu=du mii-rai ui=na?*  
 star=NOM=FOC look-POT CONT.NPST=Q  
 ‘Can you see *the stars*?’  
 b. *nau=nu=du mii-rai ui=ga?*  
 what=NOM=FOC look-POT CONT.NPST=Q  
 ‘What can you see?’ (Hayashi 2010b: 173)

The Hirara dialect of Miyako has yet another distributional pattern. According to Koloskova & Toshio (2008: 620), there is a distinction between three focus markers, namely =*ga* in content questions, =*nu* in polar questions, and =*du* in declaratives. In Ikema a special question marker for topic questions is =*da*, which is always combined with the topic marker. In Masana (Okinoerabu) the question marker =*do* can also be combined with the topic marker =*wa* (van der Lubbe & Tokunaga 2015: 362).

- (143) Ikema  
*vva=a=da?*  
 2SG=TOP=Q  
 ‘How about you?’ (Hayashi 2010b: 173, fn. 16)

Questions in **Ōgami** (spoken on Ōgami next to Miyako and in one village on Miyako itself) have a pitch that “is high and level and falls sharply on the last syllable” (Pellard 2010: 146). Similar patterns may exist for other Ryūkyūan languages but usually were not stated as clearly. There are two optional question markers, a by now familiar particle =*ka* and a suffix =*εε* that “is limited to past tense forms, the copula and stative verbs” (Pellard 2010: 151). It may be worth noting that it is identical to a suffix that derives agent nouns (Pellard 2009: 118) and we might be dealing with a development parallel to Japanese *no*.<sup>15</sup>

<sup>15</sup>For the following examples only Pellard (2009) in French was quoted, but they can usually also be found in Pellard (2010) in English.

- (144) Ōgami  
*nauripa=tu kuu-tatar-εε?*  
 why=FOC come-PST.NEG-Q  
 ‘Why didn’t you come?’ (Pellard 2009: 211)

I was unable to find a good example for the sentence-final particle =*ka* in Pellard (2009; 2010). The only example is an embedded content question.

- (145) Ōgami  
*[nau=iu as-sipa=tu tau-kaw=ka] ss-ai-n?*  
 what=ACC do-CIRC=FOC good-V=Q know-POT-NEG  
 ‘I don’t know [what I should do].’ (Pellard 2009: 225)

There is a special marker *mukara* for embedded polar questions comparable to English *if/whether* or German *ob* (Pellard 2009: 221). The focus marker =*tu* is sometimes found attached to a verb as well and we might be dealing with a development of a question marker as in Irabu, but Pellard (2009: 192) is not very clear about this.

- (146) Ōgami  
*vva=a pssnii=pa asi=tu?*  
 2SG=TOP siesta=TOP.OBJ do=?Q  
 ‘Have you taken a siesta?’ (Pellard 2009: 221)

Questions in the language **Irabu** (spoken on Irabu among the Miyako islands) exhibit an interesting interaction with focus marking. According to Shimoji (2011a: 118),

when a focus marker is present, a question marker is optional, and its form is identical to that of the focus clitic in the same clause. I treat these two (i.e., the focus marker and question marker) as different morphemes owing to the fact that they show different allomorphic patterns, even though the focus marker may be the historical source of the question marker.

If only a question marker is present, it attaches sentence-finally to the verb. This is a plain polar question.

- (147) Irabu  
*vva=a uri=u aŷ-tar=ru?*  
 2SG=TOP that=ACC say-PST=Q  
 ‘Did you say that?’ (Shimoji 2011a: 119)

In the following two examples both focus and question markers appear. The first example is a focus question, the second a content question.

(148) Irabu

a. *vva=ga=ru uri=u aŷ-tar=ru?*  
 2SG=NOM=FOC that=ACC say-PST=Q

‘Did you say that?’

b. *vva=a nau=ju=ga aŷ-tar=ga?*  
 2SG=TOP what=ACC=FOC say-PST=Q

‘What did you say?’ (Shimoji 2011a: 118)

The two markers *=ru* and *=ga* are probably cognate with Yuwan *=du* and *=ga*, where they express only focus. The fact that the question markers are optional if the focus marker is present, might be a hint of the historical development. Presumably, the focus marker *=ru* was reanalyzed as a question marker in focus questions and subsequently also marked polar questions. From there it may have spread back to focus questions in its new position attached to verbs. But in the absence of any historical data, this scenario must remain speculative. Shimoji (2011a) has one example of an embedded alternative question that shows double marking and no disjunction.

(149) Irabu

[*ssibara=ru a-tar=ru maibara=ru a-tar=ru*] *mmja s-sa-n-Ø=suga*  
 back=FOC COP-PST=Q front=FOC COP-PST=Q INTJ know-THM-NEG-NPST=but

‘But I’m not sure [whether (the house) was behind or in front].’ (Shimoji 2011a: 132f.)

The presence of the focus marker in Irabu excludes realis marking on the verb (see below on *kakari musubi*). Lawrence (2012: 396) briefly mentions question marking in the *Nakachi* variety of Miyako, which shows a somehow reminiscent pattern. In content questions there is only the marker *-ga* on the interrogative itself while polar questions have the focus marker *-ru*, exclusively. In polar questions the slightly different *-ro* is found sentence-finally.

*Hateruma* is the name of one of the Yaeyama islands but as usual is also used to refer to the language spoken there Aso (2010a; 2015). Due, however, to relatively recent population movements, the language is also spoken on another Yaeyama island, namely Ishigaki. *Hateruma* has four inferential suffixes *=kaja*, *=sa*, *=dore*, and *=paci*, the first three of which may correspond to the forms found in Okinoerabu above, i.e. *=kaja*, *=sa*, *=do* (Matayoshi 2010a: 208). But polar questions are also expressed with the enclitic *=naa* while content questions remain unmarked. This is a rather untypical pattern for a Japonic language but is the norm in most other languages in Northeast Asia (see Chapter 6).

(150) *Hateruma*

a. *da=Ø sinsin=naa?*  
 2SG=(CORE) teacher=Q

‘Are you a teacher?’

5 Survey of the grammars of questions in Northeast Asia

- b. *kuri=Ø=ja nu ja-Ø?*  
 this=(CORE)=TOP what COP-NPST  
 ‘What is this?’ (Matayoshi 2010a: 210)

Often an enclitic such as *=ba* is found in content questions, but this has instead an emphatic or focus function.

**Hatoma** is another Yaeyama variety. While Hateruma is a small island south of the main island Iriomote, Hatoma is an even smaller island on the north of it (Matayoshi 2010a: 189). Hatoma exhibits an interesting split between past and non-past content questions (Lawrence 2012: 396), the former, like polar questions, being marked by (probably rising) intonation alone and the latter showing a second split. Non-past content questions usually have an attributive form of a verb followed by the marker *-wa*. But if an interrogative phrase stands sentence-finally, it takes the marker *-ja*, instead. Apparently, the difference lies in the clause type with either a verbal or a non-verbal predicate. Content questions thus have three different markings.

(151) Hatoma

- a. *nunti kanan=wa?*  
 why write.NEG=Q  
 ‘Why won’t you write?’
- b. *kjuu=ja nuu-nu pii=ja?*  
 today=TOP what-GEN day=Q  
 ‘What day is it today?’
- c. *nuusi nat-taa?*  
 how become-PST  
 ‘What happened?’
- d. *waa aca-N k-ii ffir-un?*  
 2SG tomorrow-also come-INF give.me-AFF  
 ‘Will you come tomorrow, too?’ (Lawrence 2012: 396)

Descriptions of Ryūkyūan languages almost never give information on other question types such as alternative questions, Lawrence (2012: 397) being an exception. Hatoma alternative questions either display simple juxtaposition or double marking with the form *=kajaa*.

(152) Hatoma

- a. *kuree turu kaburee?*  
 this.TOP bird bat  
 ‘Is this a bird or a bat?’
- b. *kuree turu=kajaa kaburee=kajaa?*  
 this.TOP bird=Q bat=Q  
 ‘(I wonder) is this a bird or a bat?’ (Lawrence 2012: 397)

Cognates of the marker =*kajaa* were already encountered in Okinoerabu and Shuri. In Hatoma it can also be found in (less direct) content questions such as *taa=kajaa?* ‘(I wonder) who (is it)?’ (Lawrence 2012: 396).

The last Yaeyama variety to be considered here is called **Miyara** or Miyaran, spoken on Ishigaki island (Izuyama 2003; Davis & Lau 2015). In Miyara both polar and focus questions may be expressed with the help of rising intonation alone. In focus questions an additional focus marker =*du* appears and triggers the loss of the indicative ending on the verb.

(153) Miyara

- a. *naoja=ja sinbun jum-u-n?*  
 PN=TOP newspaper read-PRS-IND  
 ‘Will/does Naoya read the newspaper?’
- b. *naoja=ja sinbun=du jum-u?*  
 PN=TOP newspaper=FOC read-PRS  
 ‘Will/does Naoya read the *newspaper*?’
- c. *naoja=ja noo=du jum-u?*  
 PN=TOP what=FOC read-PRS  
 ‘What does/will Naoya read?’ (Davis & Lau 2015: 260)

Content questions have the same (optional) focus marker but exhibit falling intonation. Notice the absence of the final -N from content questions even if the focus marker =*du* is not present.

(154) Miyara

- zīma=ge har-u?*  
 where=DIR go-PRS  
 ‘Where are you going?’ (Davis & Lau 2015: 261)

Miyara also has the dubitative particle *kajaa* as well as a particle *i* that “indicates a request for agreement” (Izuyama 2003: 28f.). Details remain unclear, but the latter might be comparable with =*ji* in Shuri.

Yonaguni is the westernmost island of the Yaeyama islands, only about 100 km off the coast of Taiwan. Here only two Yonaguni dialects will be addressed, Dunan and Sonai. In *Dunan* polar and focus questions are marked with a sentence-final clitic =*na*. Content questions have their own sentence-final marker =*nga*. There is an additional focus marker in focus (=du) and content questions (=ba). A non verbal content question has the question marker =*ja* instead of =*nga*.

(155) Dunan

- a. *khuruma mut-i bu=na?*  
 car hold-MED IPFV=Q  
 ‘Do (you) have a car?’

5 Survey of the grammars of questions in Northeast Asia

- b. *suuti khat-i=du bu-ru=na?*  
 book write-MED=FOC IPFV-PTCP=Q  
 ‘(Are you) writing a book?’
- c. *nda=ja tharu=nki nu(=ba) thura=nga?*  
 2SG=TOP PN=DIR what(=FOC) give=Q  
 ‘What will you give to Taro?’
- d. *nma=ŋa(=ba) dunan-ccima=ja?*  
 where=NOM(=FOC) PN-island=Q  
 ‘Where exactly is Yonaguni island?’ (Yamada et al. 2015: 468, 466, 469)

Whether =*ja* might be connected to Okinoerabu and Ura =*joo* remains unclear to me.

In **Sonai** the situation is very similar to Dunan (Izuyama 2012: 442ff.). The polar question marker has the form =*na(i)* and content questions have two different markers with the same distribution, =*ga* in verbal and =*ja(a)* in non-verbal clauses. In addition, there is a dubitative form =*kaja(a)* roughly meaning ‘I wonder’ as in Hatoma and other varieties. The two elements *-du* and *-ba* obviously correspond to Dunan =*du* and =*ba*. Izuyama (2012: 443) calls them focus and selective particles but writes them attached to the preceding word with the help of a hyphen. The question markers on the other hand were written detached from the preceding word. I reanalyze all of them as enclitics.

(156) Sonai

- a. *num-i=na?*  
 drink-IND=Q  
 ‘Have you drunk it?’
- b. *nma-nki h-jun=ga?*  
 where-ALL go-PFV=Q  
 ‘Where has (she) gone?’
- c. *u=ja nu=ja?*  
 this=TOP what=Q  
 ‘What is this?’
- d. *nu=ba Nda munu=ja?*  
 what-SEL 2SG.GEN thing=Q  
 ‘Which is yours?’
- e. *nda ici=ba s-un=ga?*  
 2SG when=SEL come-PFV=Q  
 ‘When did you come?’
- f. *ta=ba=du tata-N=ga?*  
 who=SEL=FOC make.stand-CONC=Q  
 ‘Whom do you make stand?’

g. *nu=ba=du ut-iru=kaja?*  
 what=SEL=FOC fall-CONC=Q

‘I wonder which one will fall down?’ (Izuyama 2012: 439, 419, 425, 444, 421)

As in Ikema, the focus marker =*du* is also found in content questions and is not restricted to focus questions as in Yuwan.

Table 5.43 summarizes the marking of questions in Japonic languages. Given the lack of information on alternative questions, these have been excluded from the summary. In general, it appears that alternative questions show the double marked type and lack a disjunction. Forms with an additional semantic component such as those translated with ‘I wonder’ are excluded from the list as well.

Most languages have different markers for polar and content questions. Ōgami and Japanese are exceptional in allowing the same marker. Apart from Hateruma and Yilan Creole all languages have content question markers. Little information is available on focus questions. In some languages such as Dunan, Ikema, and Japanese they have the same marking as polar questions, plus an additional focus marker. In Yuwan and Shuri there are special question markers, but Shuri also allows the question and focus markers from content questions to enter focus questions. The only languages without at least an optional polar question marker are Hatoma and Miyara.

A typologically rare phenomenon of Japonic languages that is relevant for interrogative constructions is a kind of *focus concord*, usually called *kakari musubi* (KM) ‘governing (and) concordance’ (cf. Shimoji 2010: 11; Shinzato & Serafim 2013). We have already encountered a special type in Yuwan above that is limited to interrogative constructions. Specifically, the focus markers =*du* in focus questions and =*ga* in content questions necessarily are followed by the verb endings *-ui* and *-u*, respectively. Usually, however, the phenomenon is not restricted to questions but can also be found in declarative sentences. More generally, *kakari musubi* can be characterized as “a syntactic agreement construction in which specific particles called *kakari joshi* (*kakari* particles, KP henceforth) correlate with particular predicate conjugational endings other than regular finite forms to end a sentence.” (Shinzato 2015: 299)

KM is attested in some Ryūkyūan languages as well as Old Japanese, but not in modern Japanese. Altogether, Japonic has five different *kakari* particles, of which we have already encountered *ka* and *ya*. In Old Okinawan only three of them have clear cognates (Table 5.44). The first three of the markers may go back to demonstratives (cf. pre-modern Japanese demonstratives *ko-*, *so-*, *ka-*, see §5.6.3). According to Shinzato (2015), *kakari musubi* is similar to an it-cleft construction, i.e. a way of marking focus. This may be the reason why the *kakari* particles are also found in focus as well as content questions. The verbal ending triggered by the KP is usually an adnominal form. Modern Ryūkyūan languages nevertheless show several deviations from this rule. In Miyara and Irabu, for instance, there is no adnominal form of the verb (Davis & Lau 2015: 257). In Miyara the presence of the focus marker leads to the loss of the indicative ending (see also example 153 above).

5 Survey of the grammars of questions in Northeast Asia

Table 5.43: The marking of polar, focus, and content questions in Japonic; whether or not a focus marker is optional is not indicated

	PQ	CQ	FQ
East OJ	ya, ka	ka	ATTR + ya, ?ka
West OJ	ya, ka	ka	ATTR + ya, ?ka
Hachijō	?kai	?-	?
Japanese	ka#, no#	ka#	ka# + wa TOP
Yilan Creole	ga#, no#	-	?
Ura	=na# ~ =nja#	=joo#	?
Yuwan	V-mi, =na# (Table 5.41)	V-u + =ga FOC	V-ui + =du FOC
Okinoerabu	=na#, IND -ŋ, -Ø + =nja#/-jee# + PST -ti instead of -ta	=joo#	?
Shuri	V-mi, V-i, V-ti, =naa# (Table 5.42)	V-ga, V-ra + =ga FOC	=ji# + =du FOC, V-ra + =ga FOC
Tsuken	V-mi, =na#	=ga#	?
Tarama	=na#	=ga# + =ga FOC	?
Ikema	=na#	=ga# + =du FOC	=na# + =du FOC
Ōgami	=ka#, ?=tu, PST, COP, STAT.V -εε	=ka#, PST, COP, STAT.V -εε + =du FOC	?
Irabu	=ru# + =ru FOC	=ga# + =ga FOC	?
Hateruma	=naa#	-	?
Hatoma	-	PST -, ATTR + =wa, =ja# (non-verbal)	?
Miyara	-	lack of IND -N + =du FOC	lack of -N + =du FOC
Dunan	=na#	=nga#, =ja# (non-verbal) + =ba FOC	=na# + =du FOC
Sonai	=na(i)#	=ga#, =ja(a)# (non-verbal) + =ba SEL, + =du FOC	?

Table 5.44: KPs in Proto-Japonic, Old Japanese, and Old Ryūkyūan according to Shinzato (2015: 306ff.)

	Proto-Japonic	Old Japanese	Old Okinawan
Group I	*kō(swo)	koso	su
	*tyō	so	do
	*ka	ka	ga
Group II	?	ya	?yi
	?	namu	-

(157) Miyara

- a. *naoja=ja sinbun jum-u-n?*  
 PN=TOP newspaper read-PRS-IND  
 ‘Naoya will read/reads the newspaper.’
- b. *naoja=ja sinbun=du jum-u?*  
 PN=TOP newspaper=FOC read-PRS  
 ‘Naoya will read/reads the *newspaper*.’ (Davis & Lau 2015: 260)

The phenomenon found in Irabu has been called *quasi-kakari musubi* (Shimoji 2011b). Instead of the obligatory presence of a certain verb ending (usually adnominal), Irabu not only excludes the presence of realis marking but allows other types of endings (including irrealis, mood-neutral etc.).

(158) Irabu

- a. *ba=a kuruma=u=du vv-tar.*  
 1SG=TOP car-ACC=FOC sell-PST  
 ‘I sold a *car*.’
- b. *\*ba=a kuruma=u=du vv-tam.*  
 1SG=TOP car-ACC=FOC sell-PST  
 ‘I sold a *car*.’ (Shimoji 2011b: 120)

Shimoji (2011b: 121) calls these two different types positive and negative concordance. For a phenomenon similar to *kakari musubi* in NEA see §5.14.2 on question marking in Yukaghiric.

### 5.6.3 Interrogatives in Japonic

Interrogatives in Japonic languages are not very well described. Most descriptions available to me simply mention one or two forms but do not dwell on their analysis, etymology, or usage. The major exception in the Western literature is Vovin (2005: 297–336). Some interrogatives such as ‘who’, ‘what’, and ‘when’ are probably of Proto-Japonic origin (Table 5.45).

These forms represent three major groups of interrogative present in Japonic languages that start with *\*t-*, *\*n-*, and *\*e-*, respectively. Japonic has neither KIN- nor K-interrogatives. The Proto-Japonic interrogative *\*ta-* ‘who’ is basically present in all Japonic languages. Written pre-modern Japanese still had *ta-re* instead of modern day *da-re* (Ashton 1904: 63). Yilan Creole has an initial liquid instead (*la-re*). In some languages the base stem is used as interrogative while other languages exhibit different suffixes. The suffix *-re* in Japanese and its equivalents in some of the other languages is probably related to the suffix found in *do-re* ‘which’ as well as the demonstratives (see below). Its meaning is somewhat unclear but it may be treated as a stem extension.

<sup>16</sup>This form is rare and probably originates in the Western dialect.

<sup>17</sup>Hachijō data taken from <http://www008.upp.so-net.ne.jp/ohwaki/hougen.htm>. (Accessed 2016-01-19.)

5 Survey of the grammars of questions in Northeast Asia

Table 5.45: Japonic interrogatives for ‘who’, ‘what’, and ‘when’; many Southern Ryūkyūan forms stem from Bentley (2008a: 298-299); EOJ = Eastern Old Japanese, WOJ = Western Old Japanese, PMJ = pre-modern Japanese, OR = Old Ryūkyūan (Vovin 2005; Kupchik 2011); the *N* stands for prenasalization; transcription of Shodon glottal stop modified

	who	what	when
EOJ	ta-re 多例, 多礼, ... ta-(N)ka 多賀, 他加, ...	?aN- 安, ... < *ani (nani 奈尔,...) <sup>16</sup>	itu ,
Hachijō <sup>17</sup>	da-re だれ, da-ga たが	ani ぁに	?
WOJ	ta-(re) 多禮, 多礼, ... ta-Nka 多賀, 他賀, ...	nani 奈爾, 那爾, ...	itu 伊都
PMJ	ta-re, (da-re), ta-ga	nani	itsu
Japanese	da-re 誰, だれ	nani (~ nan) 何, なに	itsu 何時, いつ
Yilan Creole	la-re	nani	?
Ura	ta-ru	nan	icu
Yuwan	ta-ri ~ ta-ru	nuu	ici
Okinoerabu	ta-ru ~ ta-ŋ	nuu	?itʃi
Shodon	tha-r(u-), thaa-ga	nu(u) ~ nu(u)	?yit(i)
OR	ta(a) たぁ, ta-ru たる, ta-ga たが	nau なお	itu ?[itsi] いつ
Shuri	taa	nuu	?itʃi
Tsuken	taa	?	?
Hirara	ta-ru ~ too	noo	itsi
Tarama	taa-ga, tau	nuu	itsi
Ikema	ta-ru	nau	?
Nagahama	ta-ru	nau	itsi
Ōgami	ta-ru ~ tau	nau	iks
Irabu	ta-ru	nau	ic
Ishigaki	ta-ru ~ taa	noo	itsi
Kohama	ta-ru	nuu	itsu
Kuroshima	ta-rī ~ taa	nuu	itʃiya
Hateruma	ta-ru ~ ta(a)	nu(u)	ici
Hatoma	taa	nuu	itsi
Miyara	ta-ru	noo	itsi
Dunan	thá	nû	?
Sonai	ta(a), ta-ŋa, takka	nu(u)	ici

The suffixes *-Nka* in Old Japanese and *-ga* in Old Ryūkyūan are said to have a possessive function (Vovin 2005: 298ff.). Hachijō *-ga*, Shodon *-ga*, Tarama *-ga*, Okinoerabu *-ŋ*, and Sonai *-ŋa* are likely of the same origin. It may be worth noting, however, that in these languages the suffix combines the function of both the genitive as well as the nominative (Izuyama 2012: 417; van der Lubbe & Tokunaga 2015: 352; Aoi 2015: 415).

The Proto-Ryūkyūan interrogative meaning ‘what’ probably had the form \**nau*. Forms such as Miyara *noo* have gone through regular sound changes, in this case \**au* > *oo* (cf. Davis & Lau 2015: 258). But the connection with Japanese *nani* or Hachijō *ani* is not completely straightforward. At least one Ryūkyūan language has a form closer to Japanese (Ura *nan*), but this may be due to contact with Japanese. To my knowledge, the best, albeit problematic, explanation has been put forward by Vovin (2005: 305–313). He reconstructs a Proto-Japonic form \**nanu*, in which the \**n-* is said to be a prefix with unclear meaning. Eastern Old Japanese, Vovin claims, has a form without the prefix, as can be seen from a comparison of WOJ *naNtō*, *naNsō* and EOJ *aNtō*, *aNse*. According to Vovin, the final *-i* might derive from a suffix *-(C)i*, the meaning of which was not given. He assumes an irregular sound change in Ryūkyūan, namely the loss of the intervocalic *n*, resulting in \**nau*. Vovin (2005: 313) also notices a similarity of his reconstruction with Austronesian \**n-anu* with an unclear prefix. Blust (2013: 310) reconstructs the Proto-Austronesian form as \**anu* ‘what’, and we will encounter the Atayal form *nanu*<sup>2</sup> ‘what’ at the end of this section. The similarity is indeed striking, but depends on whether Proto-Japonic \**nanu* is a correct reconstruction or not.

However, Vovin’s explanation does not seem very plausible. For example, instead of postulating an otherwise unknown prefix *n-*, it is much more likely that Eastern Old Japanese simply lost the initial nasal that is present in Ryūkyūan as well. Let us first consider the Japanese forms *naze* and *nado* meaning ‘why’. According to Vovin (2005: 333) they have the form *naNsō* and *naNtō* in (Western) Old Japanese and are combinations of *nani* with the two defective verbs *tō* ‘to say’ and *sō* ~ *se* ‘to do’ (or a particle *sō*). Given the strong connection of the categories of REASON and ACTION, this seems plausible. Vovin (2005: 336) claims that Ryūkyūan has no cognates of the two forms, and indeed, of the references used in §5.6.2 only Shimoji (2011a: 106) mentions the two forms *nausi* ‘how’ and *nautti* ‘why’ for Irabu. Bentley (2008a: 268, 298f.) gives some additional forms (e.g., Hirara *noofii* ‘how’, *nooti* ‘why’) and reconstructs Southern Ryūkyūan (Sakishima) \**naWo-se* ‘how’ and \**naWo-nVte-* ‘why’. The *W* stands for a somewhat unclear semi-vowel \**j* or \**w* (Bentley 2008a: 218f.). Apart from certain innovations and additional suffixes found in some languages, there certainly are cognates of the Japanese interrogatives. Ryūkyūan forms such as Irabu *nau-si* suggest a derivation that is directly based on *nau* ‘what’ and the same may be true for the Old Japanese equivalents, i.e. they might be derived from \**nanu* instead of *nani*. The nasal found in some forms such as Yonaguni *nundi*, according to Bentley, was part of the suffix instead of the stem (also cf. Shuri *nuuntfi* ‘why’, Miyara 2015: 387).

The interrogative ‘when’ can be reconstructed as \**etu* (Vovin 2005: 330, see Pellard 2008: 143, passim for details on vowels). The interrogative can be found in all Japonic languages for which sufficient material is available. The analysis of PJ \**etu* is an open question but it can be classified with several other interrogatives with the resonance \**e~* > *i~* (Table 5.46). WOJ in addition has the forms *iNtu-ti* ‘where’ as well as *iku-Nta* ‘how many/much’ and EOJ *iNtu-si* ‘which’.

Several scholars have compared the interrogatives in \**e~* with Koreanic \**e-* (e.g., Frellesvig & Whitman 2004: 289; Vovin 2005: 319, 322; §5.7.3). However, a comparison based on one vowel must be treated with caution.

5 Survey of the grammars of questions in Northeast Asia

Table 5.46: Interrogatives in Proto-Japonic, Western Old Japanese (WOJ), Eastern Old Japanese (EOJ), pre-modern Japanese (PMJ), Japanese (J), and Proto-Ryūkyūan (PR) starting with *i* ~ < \**e*~ according to Aston (1904: 63ff.), Vovin (2005: 297ff.), and Kupchik (2011: 589ff.); partly modified transcription; the N stands for prenasalization

	PJ	WOJ	EOJ	PMJ	J	PR
which	*entu-re	iNtu-re	itu-re	idzu-re	do-re	*edu-re
when	*etu	itu	itu	itsu	itsu	*etu
how	*eka	ika	ika	ika	ika	*eka
how	*eku	iku-	?	iku-	iku	*eku
many						

The Old Japanese interrogative *ika* ‘how’ is not very common, is usually limited to Western Old Japanese and is followed by one of the defective copulas *n-* and *tō-* or the still more productive *nar-*, which is a contraction of *n-i ar-* ‘COP-INF exist-’ (Vovin 2005: 313–319; Kupchik 2011: 593f.). Among the cognates in Ryūkyūan languages we find Old Ryūkyūan forms such as *ika* いか, *ikyā* いきや ~ *ka* か, *kya* きや etc. and Shuri *icaa* ~ *caa* (see Vovin 2005: 318 for a more exhaustive list). In both cases there are forms with and without the initial vowel that is responsible for the palatalization of the following velar consonant. Vovin’s (2005: 317) problematic and somehow unclear conclusion is that the interrogative has to be analyzed as \**e-ka*. But this is no explanation for why the initial element—which must be considered the interrogative as such—can simply be omitted. It is more reasonable to assume that *ika* was considered an inseparable interrogative by the speakers, which is why the, maybe irregular, loss of the vowel did not affect its interrogative status as such. The same criticism also applies to his explanation of the other interrogatives that will be addressed in the following. Japanese *ikaga* ‘how’ derives from the Old Japanese fixed expression *ika n-i ka* ‘how COP-INF Q’ (Vovin 2005: 314, fn. 120). Vovin (2005: 319) compares the hypothetical element *-ka* with Korean but leaves open any further detail.

Apart from the locative endings, the Old Japanese interrogative *iNtu-ku* ‘where’ has a direct cognate in Old Ryūkyūan *idu-ma* > *zuma* ずま as well as in modern Ryūkyūan languages such as Miyara *zima* (Vovin 2005: 321; Davis & Lau 2015: 261). The second part *-ma* is claimed to be a noun meaning ‘place’, but in this case the interrogative *idu-* would be expected to have the meaning ‘what’ or ‘which’ rather than ‘where’. In fact, from a typological perspective PJ \**entu* (together with the extended form \**entu-re*) likely was a selective interrogative ‘which’ at first and only later developed into a locative interrogative ‘where’, as it was combined with a locative marker or a noun meaning ‘place’ (*-ma*). Several languages of the region have parallel developments and this scenario is corroborated by data from some Ryūkyūan languages such as Irabu *nzi* ‘which’ versus *nza* ‘where’ that may go back to the plain and derived forms, respectively. Ōgami still has the non-palatalized forms *nti* (~ *iti*) versus *nta* (~ *ita*) that make this development seem

more plausible. However, Ryūkyūan languages show much stronger variation in forms meaning ‘where’ than in those interrogatives previously encountered. Some of Vovin’s (2005: 321, especially fn. 123) otherwise good explanations for those deviations are somewhat speculative and cannot be taken at face value. Among the dialects mentioned in §5.6.2, for example, we find the forms listed in Table 5.47. A possible explanation for Shuri *maa* is the loss of the first part of *idu-ma*. All other forms can, following Vovin, be derived directly from *idu-ma* or rather its predecessor PR \**eNtuma* (Vovin 2005: 321, fn 123). But this is certainly not true for Yuwan *daa*, in which the first part was deleted as well (cf. Okinoerabu *?uda*).

Table 5.47: Interrogative forms meaning ‘where’ in Japonic

Language	Form
Eastern Old Japanese	iNtu- 伊豆
Western Old Japanese	iNtu-ku 伊豆久
(written) pre-modern Japanese	idzu-ko
Hachijō	do-ko
Japanese	do-ko 何処
Yilan Creole	do-ko
Ura	?uda
Yuwan	daa
Okinoerabu	?uda
OR	idu-ma > zuma すま
Shuri	maa
Tsuken	maa
Ōgami	nta ~ ita
Irabu	nza
Hateruma	za
Miyara	zima
Dunan	nmâ
Sonai	nma

Vovin mentions a Northern Ryūkyūan form *raa*, not encountered thus far, that is probably a variant of *daa*. The distinction between LOCATION, DIRECTION, and SOURCE has not been given for the majority of languages. Most likely, the difference in most languages is indicated with case markers as in (Eastern) Old Japanese (*iNtu-yu* ‘where from’), Japanese (*doko ni* ‘where (to)’, *doko e* ‘where to’, *doko kara* ‘where from’, my knowledge), or Ura (*?uda=ne* ‘where’, *?uda-gatfi* ‘where to’, van der Lubbe & Tokunaga 2015: 361).

In modern Japanese only a few forms in *i*- survive (e.g., *itsu*, *ikura*), which is due to a replacement with forms built on the stem *do*-. The fact that all forms are analyzable shows that this is a relatively new system. In fact, the interrogative stem *do*- in Japanese is completely in line with the demonstratives (Table 5.48). These paradigms were clearly

5 Survey of the grammars of questions in Northeast Asia

at least partly present in Old Japanese (Table 5.49). But in standard Japanese the distal demonstrative *ka-* has been replaced with *a-* and Old Japanese still lacked the stem *do-*. Interestingly, written pre-modern Japanese still had forms based on the stems *ka-* and *idzu-* (Table 5.50). In Japanese the word *kare* started out as a demonstrative, changed its meaning to a male third person pronoun and also means ‘boyfriend’ today.

Table 5.48: Parallels in demonstratives and interrogatives in Japanese (based on Dixon 2012: 407; Hasegawa 2015: 332); the Kansai dialect has a regular form *a-ko* instead of the irregular *a.so-ko*; some endings were omitted

	proximal	medial	distal	interrogative
pronominal	ko-re	so-re	a-re	do-re
adnominal	ko-no	so-no	a-no	do-no
place	ko-ko	so-ko	a.so-ko	do-ko
thing/person (vulgar)	ko-itsu	so-itsu	a-itsu	do-itsu
direction/person (polite)	ko-chira	so-chira	a-chira	do-chira
type/kind	ko-nna	so-nna	a-nna	do-nna
adverb	ko-o	so-o	a-a	do-o

Table 5.49: Old Japanese demonstrative and interrogative paradigms (Vovin 2005: 272; Kupchik 2011: 583, partly modified); there are additional forms such as *wote* ‘that (over there)’ not shown here

	proximal	medial	distal	interrogative
pronominal	kō-(re)	sō-(re)	ka-(re)	WOJ iNtu-re
adnominal	kō-nō	sō-nō	ka-nō	?
place	kō-kō	sō-kō	-	WOJ iNtu-ku

Table 5.50: Paradigms of written pre-modern Japanese demonstrative and interrogative paradigms (Aston 1904: 60ff.)

	proximal	medial	distal	interrogative
pronominal	ko-(re)	so-(re)	ka-(re)	idzu-(re)
adnominal	ko-no	so-no	ka-no	idzu-re-no
place	ko-ko	?so-ko	?	idzu-ko

This paradigmatic parallel between pre-modern *idzu-* and modern *do-* might suggest that it is in fact the same etymological entity in a different phonological shape. In some Ryūkyūan languages there is a form without the initial vowel as well. For example, Okinoerabu *ʔuduru* ‘which’ and *ʔuda* ‘where’ (van der Lubbe & Tokunaga 2015: 350) must

Table 5.51: Paradigms of Hachijō demonstrative and interrogative paradigms (Kokuritsu Kokugo Kenkyūjo 1950: 204f.); cf. *dare* ‘who’; several dialectal forms were omitted

	proximal	medial	distal	interrogative
pronominal	ko-re	so-re	u-re	do-re
adnominal	ko-no	so-no	u-no	do-no
place	ko-ko	so-ko	u-ko	do-ko

directly correspond to *diru* and *daa* in Yuwan. The paradigms in Hachijō are very similar to modern Japanese, but there is a different distal stem *u-* that looks similar to the medial stem in Ryūkyūan (Table 5.51). In general, the Northern Ryūkyūan languages, especially Amami Ryūkyūan languages, have a pattern very similar to Japanese. Except for Miyara, the Southern Ryūkyūan languages do not exhibit the same similarities in demonstrative and interrogative paradigms. Table 5.52 to Table 5.56 show paradigms for those languages that were described in sufficient detail. Also, northern Ryūkyūan shares the distal stem *a-* with modern Japanese, while southern Ryūkyūan still has *ka-*, as does Old Japanese. What is more, the extension of the demonstrative and the interrogative are only found in northern Ryūkyūan and are not necessarily identical in form. In Yuwan and Shuri, for example, the demonstratives have the extension *-ri* ~ *-ri*, but the demonstrative has *-ru*. In Dunan, the extension can only be found in the distal demonstrative.

Apparently, instead of the selective interrogative, Yonaguni uses an objective inter-

Table 5.52: Paradigms of Yuwan (Amami) demonstrative and interrogative paradigms (Niinaga 2010: 50f.); cf. *ta-ri/ru* ‘who’; see also Martin (1970: 123-124)

	proximal	medial	distal	interrogative
pronominal	ku-ri	u-ri	a-ri	di-ru
adnominal	ku-n	u-n	a-n	di-n
place	ku-ma	u-ma	a-ma	daa

Table 5.53: Paradigms of Shuri (Okinawan) demonstrative and interrogative paradigms (Miyara 2015: 387); form in square brackets from OCLS (1999/2003)

	proximal	medial	distal	interrogative
pronominal	ku-ri	?u-ri	?a-ri	dzi-ru
adnominal	ku-nu	?u-nu	?a-nu	[dzi-nu]
place	ku-ma	?u-ma	?a-ma	maa

5 Survey of the grammars of questions in Northeast Asia

Table 5.54: Paradigms of Ōgami (Miyako) demonstrative and interrogative paradigms (Pellard 2009: 123; 2010: 129), cf. *ta-ru* ‘who’; no forms with the Ōgami adnominal (genitive) *-nu* are available; gaps are filled with forms from Miyako proper in square brackets (OCLS 1999/2003)

	proximal	medial	distal	interrogative
pronominal	ku-ri	u-ri	ka-ri	nti ~ iti
adnominal	[ku-nu]	?	[ka-nu]	[nza-nu]
place	?ku-ma	u-ma	ka-ma	nta ~ ita

Table 5.55: Paradigms of Miyara (Yaeyama) demonstrative and interrogative paradigms (Izuyama 2003: 24), cf. *ta-ru* ‘who’; there are also the forms *nge ~ nga* ‘there (medial)’ and *zINGe ~ zINGa* ‘where’ (+ *-ge ~ -ga*)

	proximal	medial	distal	interrogative
pronominal	ku-ri	u-ri	ka-ri	zi-ri
adnominal	ku-nu	u-nu	ka-nu	?
place	ku-ma	u-ma	ka-ma	zi-ma

rogative, e.g. Sonai *nu* ‘what’ *nu-nu* ‘what-ADJ’ (Izuyama 2012: 431).

Less complicated than the locative forms are the quantitative interrogatives ‘how much’ and ‘how many’ that are based on PJ *\*eku*. Two suffixes, *-Nta* (maybe a collective) and *-ra* (maybe a plural) can sometimes be found attached to the stem (Vovin 2005: 330, fn. 129). Whether *\*eku* was analyzable or not remains an open question. Middle Japanese had another variant *iku-tu* ‘how many’ that is not attested in Old Japanese. Ryūkyūan languages have cognates of Old Japanese *\*eku* and *\*ekura* as well as of Middle Japanese *ikutu*. Similar to *\*eka* the initial vowel was sometimes lost and in some cases led to the palatalization of the following velar, e.g. Benoki *kassaa* (Vovin 2005: 332), but Yuwan *ik-jassa* (Niinaga 2010: 51) < *iku-ra* ‘how much’. In some languages the interrogative *\*eku* is preserved and is usually combined with a classifier, e.g. Okinoerabu *?iku-tfi* ‘how many things’, *?iku-tai* ‘how many people’ (van der Lubbe & Tokunaga 2015: 351). In Japanese

Table 5.56: Paradigms of Dunan (Yonaguni) demonstrative and interrogative paradigms (Yamada et al. 2015: 454, 456f.)

	proximal	medial	distal	interrogative
pronominal	khú	ú	kha-ri	?
adnominal	khu-nu	u-nu	kha-nu	?
place	khú-ma	û-ma	khá-ma	nmâ

this pattern has been taken over by *nan-* followed by a classifier, e.g. *nan-mei* 何名 ‘how many people’ (which is the source of Yilan *name*, Peng Qiu 2015: 53).

Table 5.57 shows those interrogatives found in written and spoken pre-modern Japanese interrogatives. Except for those forms based on *idzu*, the interrogatives are still present in modern Japanese. There are the resonances *i-* and *n-*. Today there is also a resonance in *d-*, but in written pre-modern Japanese, the interrogative *tare* ‘who’ was unique in that it did not exhibit any of the resonances. Japanese *dare* with an initial *d* might be an innovation based on *dore*.

Table 5.57: Pre-modern Japanese interrogatives (Aston 1904: 63ff.); forms marked with an asterisk \* are limited to the written language; not all derivations are shown

Meaning	Form
who	ta-re*, da-re
what	nani
what-PL	nani-ra
adnominal	nani-no (> nanno)
who (Japanese <i>hito</i> ‘man’)	nani-bitō
why	naze
why	nado
which	idzu-re*, do-re
which (adnominal)	idzu-re-no*, do-no
where	idzu-ko*, do-ko
whither	idzu-chi*, do-chi
whither	idzu-kata*
how, what manner	ika
how many	iku-tsu
how much	iku-ra
when	itsu

Few descriptions of Ryūkyūan languages available to me give such an exhaustive list of interrogatives. Some questions are thus hard to answer. But the limited data allow the observation that, from a typological point of view, the interrogative systems are very different from one another. In Hateruma, for instance, all attested interrogatives except *ici* ‘when’ are only two phonemes long and none is readily analyzable synchronically (*nu* ‘what’, *za* ‘where’, *ne* ‘why, how’, *ta* ‘who’, Matayoshi 2010a: 199; 2015: 429).

5 Survey of the grammars of questions in Northeast Asia

Table 5.58: Ōgami interrogatives (Pellard 2009: 132; 2010: 129); my tentative analysis based on Pellard (2009; 2010)

Meaning	Form	Analysis
who	ta-ru ~ tau	< PJ, loss of <i>r</i> , analogy to <i>nau</i> ?
when	iks	< PJ
how many	if- + CLF	e.g., if-tau ‘how many people’
what	nau	< PR
which	nti ~ iti	
where	nta ~ ita	analogy to <i>nti</i> ~ <i>iti</i>
why	nau-ripa	circumstantial converb - <i>ripa</i>
how	nau-pasi	
how much	nau-nu-pusa, nti-ka-pusa	

In Ōgami, on the other hand, the interrogatives are up to nine phonemes long and some are at least partly analyzable (Table 5.58). Ōgami has two main resonances *i*- and *n*- as well as one form *taru* ‘who’ that does not partake in any of them. The two Ōgami forms inquiring about quantity apparently are based on *nau* ‘what’ and *nti* ‘which’, respectively, and can be analyzed as *nau-nu-pusa* ~ *nti-ka-pusa*. The exact meaning of the suffixes remains unclear, however. A connection to the desiderative form *-pus* is unlikely on semantic grounds. The second part of *nau-pasi* also remains unclear. There is a circumstantial converb form *-ripa* (Pellard 2009: 146) that might have been attached to a hypothetical interrogative verb *nau-* ‘to do what’, yielding *nau-ripa* ‘why’.

In Yuwan *nuusjattu* probably has a similar background and may be an amalgamated form containing the elements *nuu* ‘what’, the verbalizer *-s(j)ar*, and the past causal converb *-tattu* (Niinaga 2010: 66, 71). Japanese *dō yatte* literally means ‘doing how’ and can be analyzed into *dō* ‘how’ and the so-called *te*-form (roughly gerund) of the verb *yaru* ‘to do, to give, to put’. These few cases suffice to show a strong connection between the two categories of ACTIVITY and REASON (§4.3).

The Amami languages **Yuwan** and **Shodon** as well as the Okinawan language **Shuri** (Table 5.59) exhibit a pattern very similar to Japanese and have the three resonances *n*-, *i*- and *d*- (> *dʒ* in Shuri). But the languages preserve an initial unvoiced aspirated plosive *t* in the interrogative meaning ‘who’.

The only polysemy that has been described can be found in Hateruma *ne*, which covers both MANNER and REASON.

For the most part, interrogatives in **Yilan Creole** are identical or almost identical to Japanese (e.g., *lare* ‘who’, *nani* ‘what’, *doko* ‘where’, *ikura* ‘how much’, *name* ‘how many (people)’, Peng Qiu 2015: 52ff.). One interesting phenomenon as opposed to Standard Japanese (107c) is the use of an interrogative basically meaning ‘who’ instead of ‘what’ in questions about names, see also (140b) from Tarama (see Idiatov 2007; Hölzl 2014b for a general discussion). This may be due to influence from Austronesian languages, maybe via Mandarin Chinese as spoken on Taiwan.

Table 5.59: Shodon (Martin 1970: 123f.), Yuwan (Niinaga 2010: 51), and Shuri (Miyara 2015: 387) interrogatives; accents removed, modified transcription for glottal stops; form in square brackets from OCLS (1999/2003)

Meaning	Shodon	Yuwan	Shuri
who	tha-r(u-), thaa-ga	ta-ri ~ ta-ru	taa
what	nu(u)	nuu	nuu
why		nuusjattu	nuuntfi
which one	dir	diru	dziru
which (adn.)	din	din	[dʒi-nu]
where	da(a)	daa	maa
what kind of	ʔyikhyassyun		
when	ʔyit(i)	ici	ʔitfi
how many	ʔyitkhut(i) ~ ʔyitkhu(u)t(i)		
how much	ʔyikhyassa	ikjassa	tʃassa
how			tʃaffi

(159) Yilan Creole

*anta no namae ga lare?*

2SG GEN name TOP who

‘What is your name?’ (Peng Qiu 2015: 54)

(160) Japanese

*anata no namae wa nan desu ka?*

2SG GEN name TOP what COP Q

‘What is your name?’ (constructed in analogy to Yilan Creole)

(161) Mandarin Chinese (Taiwan)

*nǐ de míngzi shì shéi?*

2SG GEN name COP who

‘What is your name?’<sup>18</sup>

(162) Maryinax Atayal

*ima<sup>2</sup> a<sup>2</sup> ralu<sup>2</sup>=su<sup>2</sup>?*

who NOM.NRF name=2SG.BG

‘What is your name?’ (Huang 1996: 271)

This might be an areal trait that has its origin in Austronesian languages where it is a rather typical phenomenon (Blust 2013: 509f.). Standard Chinese as spoken in the People’s Republic of China usually employ the interrogative *shénme* ‘what’. Other varieties

<sup>18</sup>This sentence was given to me by a native speaker from Taiwan during my talk at *The 8th International Conference on Construction Grammar* (Hölzl 2014b). Chinese also has further constructions.

of Atayal such as Wulai in turn employ *nanu* ‘what’ instead of *ima* ‘who’ in official contexts, which may have its origin in Chinese (Huang 1996: 293). While in Yilan Creole the use of *lare* ‘who’ may have an origin in Austronesian, the whole construction rather resembles Chinese and especially Japanese, except for the lack of the copula.

## 5.7 Koreanic

### 5.7.1 Classification of Koreanic

Korean has a North Korean (Pyongyang) and a South Korean (Seoul) standard. Here primarily the latter will be considered. In addition, Korean is officially recognized as a minority language in China, where it has developed its own standardized version of Korean based on the language spoken in Yanbian, Jilin province (L. Brown & Yeon 2015: 466). But apart from the standard languages, Korean also contains a considerable amount of dialectal variation. Usually, six different dialect areas are recognized (L. Brown & Yeon 2015: 461), but it has become increasingly clear that Yukcin has to be considered a seventh dialect (e.g., King 2006b: 130).



Sohn (1999: 58) also differentiates between seven dialect zones, but instead of Yukcin he regards Chungcheong, included in the Central Dialect above, as a separate entity. Jeju clearly is the most aberrant member of the Korean dialects (e.g., Kiaer 2014). Vovin (2013b) even goes so far as to consider Jeju a Koreanic language in its own right. He claims that the primary division is between Jeju on the one side and the varieties spoken on the Korean Peninsula on the other. In his view, Yukcin, part of the Northeastern dialect area, is also sufficiently different from the rest of the dialects to consider it a separate language. But Sean (2015: 8) recently came to the rather convincing conclusion “that the early historical relationships among Koreanic variants are considerably non-treelike”. In general, it may thus be better to conceptualize Koreanic as a dialect continuum with strong mutual contacts that make a classification into different languages problematic.

Within the Northeast Asian area, apart from the Korean Peninsula and adjacent regions in China, significant numbers of Korean speakers can also be found on Sakhalin, in Japan, and in Central Asia. The language in Central Asia, mostly in Uzbekistan and Kazakhstan, has its origin in Northeastern and Yukcin dialects, while the language spoken on Sakhalin is ultimately derived from the Southeast of Korea (King 2006b: 128). It is primarily the language spoken in Central Asia—also known as Kolyemal (Koryo language)—that will be included in this chapter. The Korean dialects in China are not very well described, but one can roughly state that “Yanbian Korean has its roots in Hamgyong dialect, whereas the variety of Korean spoken in Liaoning is of the Pyongan variety and that of Heilongj[i]ang is based on Gyeongsang” (L. Brown & Yeon 2015: 466, corrected). Given the scarcity of resources, only the variety spoken in Yanbian, Jilin province, will be included in this study (Zhao Xi 1982; Xuan Dewu et al. 1985). In Japan, apart from mainland Korean dialects, we also find speakers of Jeju, especially in isiŌsaka (Saltzman 2014).

### 5.7.2 Question marking in Koreanic

When it comes to question marking, Korean has a complicated split system that depends on the speech level. The interrogative forms in Korean qualify as interrogative mood markers because they are in complementary distribution with declarative markers. In other words, the interrogative suffixes replace the declarative ones and are not merely attached to them. This is a major difference compared to most languages in Northeast Asia.

(164) Korean (Jilin)

- a. *narssi-ka tfoh-ta.*  
weather-NOM good-DECL.PLAIN  
'The weather is good.'
- b. *narssi-ka tfoh-ni?*  
weather-NOM good-Q.PLAIN  
'Is the weather good?' (Xuan Dewu et al. 1985: 57)

Descriptions disagree in the number of forms and speech levels in Korean. Table 5.60 shows these according to the analysis by Song (2005), who distinguishes six different levels. There are declarative, interrogative, imperative, and propositive endings. The suffixes are usually called “sentence enders”, because they always take the last position in a sentence and are not restricted to verbs as such, but can also attach to verbal adjectives. Consider the following examples from Jilin Korean.

(165) Korean (Jilin)

- a. *ka-nunka?*  
go-Q.FAM  
'Are (you) going?'

- b. *k'w-nka?*  
 big-Q.FAM  
 'Is (it) big?' (Xuan Dewu et al. 1985: 31)

Table 5.60: Korean sentence enders (Song 2005: 125)

	Statements	Questions	Commands	Proposals
Plain	-(n)ta	-ni/-(nu)nya	-ela/-ala	-ca
Intimate	-e/-a	-e/-a	-e/-a	-e/-a
Familiar	-ney	-na/-nunka	-key	-sey
Semi-formal	-o	-o	-(u)o	-(u)psita
Polite	-eyo/-ayo	-eyo/-ayo	-eyo/-ayo	-eyo/-ayo
Deferential	-(su)pnita	-(su)pnikka	-(u)sipsio	-(u)sipsita

Some of the sentence enders can be further analyzed. The first element in *-n-unya*, *-n-un-ka*, and maybe in *-n-i*, as well as the medial element in *-sup-ni-kka* may be an indicative marker. The suffix *-sup* is an addressee honorific while the suffix *-un* has been called a “pre-nominal-modifier” suffix. The polite forms are identical with the intimate forms except for an additional suffix *-yo* (Sohn 1994: 337ff.). In the Chungcheong dialect, often included into the Central dialect, but treated as a separate dialect by Sohn (1999: 58), this takes the characteristic form *-yu* (L. Brown & Yeon 2015: 462).

Some of the forms in Table 5.60 are not restricted to one function. In fact, of the interrogative forms mentioned, only the plain, familiar, and deferential forms are not also found in statements, commands, or proposals. Sohn (2015: 449) lists additional variants for plain statements (*-la* instead of *-ta*) and semi-formal (*-(s)o/-(s)wu* instead of only *-o*) questions.

The sentence endings in the officially recognized variety of Korean spoken in China are very similar to standard Korean (Table 5.61). The authors mention additional forms not shown here such as *-tʃi* or *-tʃio*, which are found in all sentence types. These probably correspond to the committal *-ci* and its combination with the polite marker *-ci-yo* > *-cyo* in Standard Korean (see below). There are, furthermore, the endings *-(num)tʃi*, *-(num)ja*, and *-najo* that are restricted to the interrogative sentence type. Their exact difference in meaning remains unclear. But these are clearly combinations of other elements already encountered. The element *-num* is known from the complex familiar interrogative ending *-nun-ka* and *-najo* is the familiar interrogative ending *-na* in combination with the suffix *-jo* known from the polite speech level. The last elements in *-nun-tʃi* and *-nun-ja* are probably the marker *-tʃi* seen before and the intimate marker *-(j)ə/-a*, respectively, both of which are speech act neutral.

As opposed to standard Korean *-(u)si-psita*, Chinese Korean *-(w)psita* lacks the element *-si* that is present in *-(u)si-psiyo/-(w)si-psiyo* and has been characterized as a “subject honorific suffix” (Sohn 1994: 344). For Standard Korean Kim-Renaud (2012: 151) mentions an additional set of so-called “superdeferentials”, the interrogative form of which

Table 5.61: Sentence enders in Korean as spoken in China (Xuan Dewu et al. 1985: 62f.; Zhao Xi 1982: 75) listed analogous to Table 5.60

	Statements	Questions	Commands	Proposals
Plain	(-nuu)-(n)ta	-ni	-(j)əra/-ara	-tʃa
Intimate	-(j)ə/-a	-(j)ə/-a	-(j)ə/-a	-(j)ə/-a
Familiar	-ne	-na/-(num)ka	-ke	-se
Semi-formal	-(s)o	-(s)o	-(s)o	-(u)psita
Polite	-(j)əjo/-ajo	-(j)əjo/-ajo	-(j)əjo/-ajo	-(j)əjo/-ajo
Deferential	-(su)pnita	-(su)pnikka	-(u)sipsiyo	-(u)psita

is *-(u)si-naikka*. According to her, the familiar interrogative forms (called “deferential equal”) are *-(n)(u)nka(yo)* and *-(u)si-na(yo)*. In the latter form, both the honorific suffix *-si* and the polite marker *-yo* are optional, and the same is true for *-(u)si-psita/-(u)psita* and other sentence enders. Variants with either the vowel *e* or *a* depend on the vowel in the preceding syllable. The variant with *a* follows syllables that contain an *a* or an *o*, otherwise the variant with *e* is employed. This is a special kind of restricted vowel harmony still present in Korean. Table 5.62 shows all attested standard Korean variants with the help of two verbs and two adjectives.

Table 5.62: Interrogative paradigms of two verbs and two adjectives in Korean (Sohn 1994: 15-16)

	mek- ‘to eat’	po- ‘to see’	coh ‘good’	si ‘sour’
Plain	mek-ni	po-ni	coh-(u)ni	si-ni si-nya
	mek-nunya	po-nunya	coh-unya	
Intimate	mek-e	po-a	coh-a	si-e
Familiar	mek-na	po-na	(coh-na)	(si-ne) si-nka
	mek-nunka	po-nunka	coh-unka	
Semi-formal	mek-so	po-o	coh-so	si-o
	mek- <b>uo</b>			
Polite	mek-eyo	po-ayo	coh-ayo	si-eyo
Deferential	mek-supnikka	po-pnikka	coh-supnikka	si-pnikka

The use of the different speech levels is highly complex and has been very well summarized by Song (2005: 126f.), whose concise description is worth quoting in an abbreviated form. See Brown (2011) for details.

The **plain** speech style is used between friends or siblings whose age difference is not substantial (perhaps a one or two year age gap; in Korean culture, a three or more year age difference is regarded as substantial), or by old speakers (e.g. parents or teachers) to young children. [...]

The **intimate** speech level is referred to as *panmal* ‘half talk’ in Korean. This level is similar to the plain level in that it is used between close friends and siblings (both before middle age), by young school children to adult family members (especially their (grand)mother but probably not their (grand)father) or by a man to his (younger) wife. [...]

The **familiar** speech level is used to someone who has a lower social status than the speaker. When this level is chosen, however, the speaker is signal[1]ing a reasonable amount of courtesy to the hearer. [...] it is typically used by male adults to younger male adults who are probably under the former’s influence (e.g. protégés or former students), or to their sons-in-law. [...]

The **semi-formal** speech-level [...] has almost completely fallen into disuse and may indeed sound old-fashioned to young people’s ears. It is definitely a speech level associated with the older generation. If used, however, it is to someone with lower social status than the speaker and it is regarded as a slightly more courteous speech level than the familiar speech level. [...]

The **polite** speech level, together with the intimate speech level, is the most commonly used speech level, but, unlike the intimate speech level – which is emblematic of intimacy, familiarity or friendliness – it is used when politeness or courtesy is called for, regardless of the social status of the hearer, as long as they are old enough (university students and older). [...]

Finally, the **deferential** speech level is the highest form of deference to the hearer. This speech level is thus used to people with unquestionable seniority. It is never used to someone with equal or inferior social status. [...] (my boldface)

However complicated the internal division of question marking may be, it does not depend on the question type. The following content questions display the same question markers as did the polar questions above. Interrogatives remain *in situ* (Sohn 1999: 265) but nevertheless are often in sentence initial position.

- (166) Korean (Jilin)  
*muyəs-ur ha-nunka?*  
 what-ACC do-Q.FAM  
 ‘What are (you) doing?’ (Xuan Dewu et al. 1985: 42)

- (167) Korean  
*mues-ul ha-ni?*  
 what-ACC do-Q.PLAIN  
 ‘What are (you) doing?’ (Song 2005: 146)<sup>19</sup>

Notice the slight dialectal differences such as the presence of an intervocalic consonant in Jilin Korean *muyəs* as opposed to standard Korean *mues* (also cf. *-(u)si-psiyo* versus

<sup>19</sup>In casual speech this sentence is said to be pronounced *mwel hani*.

-(u)si-*psio*), as well as the difference in speech level. Alternative questions do not exhibit an obligatory disjunction. Instead, each alternative takes one of the interrogative sentence enders listed above. Naturally, the two markers have to be identical, i.e. are from the same speech level.

- (168) Korean (Jilin)  
*kitʃa-ka məntʃə o-r-ka, tʃatoŋtʃa-ka o-r-ka?*  
 train-NOM first arrive-PRS-Q.FAM car-NOM arrive-PRS-Q.FAM  
 ‘Does the train or does the car arrive first?’ (Xuan Dewu et al. 1985: 94)

- (169) Korean  
*wuli-ka ka-l-kka.yo salam-ul ponay-l-kka.yo?*  
 we-NOM go-PRS-Q.FAM person-ACC send-PRS-Q.FAM  
 ‘Shall we go or shall (we) send someone?’ (Sohn 1994: 122)

In the latter example the same politeness marker *-yo* that we have already encountered in the polite level endings *-a.yo ~ -e.yo* is found in the Standard Korean example.

There is an optional disjunction *an-i-myen* ‘NEG-COP-COND’ that literally means ‘and if not’ (Sohn 1994: 20) and is thus a parallel to Mongolian *eswel* (§5.8.2).

- (170) Korean  
*yongho-ka te khu-ni, animyen nami-ka te khu-ni?*  
 PN-NOM more big-Q.PLAIN or PN-NOM more big-Q.PLAIN  
 ‘Is Yongho taller or Nami?’ (Sohn 1994: 20)

Negative alternative questions may make use of a negative verb such as in the idiomatic expression in (171).

- (171) Korean  
*ka-l-kka ma-l-kka?*  
 go-PRS-Q NEG-PRS-Q  
 ‘whether to go or not’ (Sohn 1999: 392)

When the first alternative is a copula, the second alternative has to be the negative counterpart of a copula.

- (172) Korean  
*canton iss-ni eps-ni?*  
 change COP-Q.PLAIN NEG-Q.PLAIN  
 ‘Do you have change or not?’ (Kim-Renaud 2012: 150)

These are constructions very similar to those of surrounding languages such as Japonic, Mongolic, or Tungusic (see Chapter 6).

Yoon (2010: 2783) investigated the relative frequency of question types. In this study there were 70% polar questions (including tag questions), 29% content questions and only

## 5 Survey of the grammars of questions in Northeast Asia

3% alternative questions. However, 15% of all polar questions were actually tag questions. The situation is thus very similar to Japanese (§5.6.2). Tag question markers usually have the form *ku-ci-yo* ‘do.so-COMM-POL’ (> *kucyo*) and are attached to a declarative sentence.

Instead of completing a statement with a sentence ending and then adding a tag question such as *ku-ci-yo*, it is possible to put *-ci* or the contracted form of its negative form *-canh* [*< -ci-anh*] into the sentence ending of the main statement without using it in a separate tag question. Such a tag question marked in the sentence ending is called a “pseudo-tag question” by some researchers (Yoon 2010: 2788, my brackets)

The author has recorded the two examples in (173):

(173) Korean

- a. *hwuchwu mac-cyo?*  
black.pepper correct-COMM.POL  
‘It is black pepper, right?’
- b. *wenlay khu-canh-a.yo?*  
originally tall-COMM.NEG-POL  
‘She has been tall since birth, right?’ (Yoon 2010: 2788)

I was unable to find information on focus questions in the literature available to me. The following examples were elicited from a native speaker in South Korea via internet in April 2016. Focus was expressed in this case with word initial position of the focused element. The analysis roughly follows Song (2005).

(174) Korean

- a. *nayil hakkyo-ey ka-pnikka?*  
tomorrow school-DIR go-Q.DEF  
‘Do (you) go to school tomorrow?’
- b. *hakkyo-ey nayil ka-pnikka?*  
school-DIR tomorrow go-Q.DEF  
‘Do (you) go to school tomorrow?’
- c. *tangsin-un nayil hakkyo ka-pnikka?*  
2SG-TOP tomorrow school go-Q.DEF  
‘Do you go to school tomorrow?’ (elicited, slightly adjusted)

Similar to Japanese, the question marker does not change its form and remains in sentence-final position. A topic marker *-(n)un* attaches to the focused pronoun in the last example that takes sentence initial position. The other sentences do not have an overt pronoun, as “Koreans tend to avoid second-person pronouns altogether” (Song 2005: 75). The second sentence differs from the first in the sentence initial position of *hakkyo-ey* (cf. Song 2005: 107).

Available descriptions of questions in Jeju are not very specific or detailed. Kiaer's (2014: 13f.) otherwise good description only gives an unanalyzed list of 19 different interrogative endings: *-ka(ko)*, *-n'ga(go)*, *-nya*, *-ne*, *-nda*, *-tia(ti)*, *-lle*, *-chi*, *-k'o/-llogo*, *-ra*, *-men*, *-sǒ*, *-an*, *-sun/-mnekka*, *-ptega*, *-ptegang*, *-sugang*, *-sukkwa(gwa)*, and *-suga(kka)*. Unfortunately, there is no information on the semantic or pragmatic differences between all these suffixes and it is doubtful that they all simply mark questions. One may only speculate that they fall within different registers that are based on politeness. The interesting examples given by Kiaer (2014) lack a morpheme analysis and a glossing, which makes their analysis rather unclear. For instance, the three sentences in (175) were all translated as 'Where are you going?'

(175) Jeju

- a. *ōdi kamdi?*
- b. *ōdüi kamini?*
- c. *ōdi kamsini?*

'Where are you going?' (Kiaer 2014: 14, 16, 17)

The interrogative *ōdi* corresponds, of course, to Korean *eti* 'where (to)' and *ka-* in both languages means 'to go'. The suffix *-m* is considered to be a marker for the present tense but is better understood as an indicative marker (e.g., Saltzman 2014: passim). The analysis of *-amsi* as a marker for progressive aspect is equally problematic. The final *-ni* might be comparable to the plain question ending in Korean. But neither *-ni* nor *-di* are listed as an interrogative ending by Kiaer, who also leaves open the difference between *ōdi* and *ōdüi* (maybe a typographic error). Sohn (1999) provides a more complete analysis of Jeju interrogative sentence enders, which is given in Table 5.63 below. Among these we find the two plain level question markers *-(e)m-ti(ya)* and *-(e)m-sini*, which correspond to *-m-di* and *-m-sini* in (175a, 175c), but no correspondence to *-mini* (175b) was found. Possibly, *-mi-ni* is the same ending as *-(e)m-si-ni*, but without the suffix *-si*. It may be noted that the expression 'Where are you going?' is a common greeting in Korean that exists on different speech levels. In this expression the marker *-si* is optional on all speech levels, which corroborates the analysis of the Jeju ending as *-mi-ni*.

(176) Korean

- a. *eti ka-si.pnikka?*  
where go-Q.DEF
- b. *eti ka-pnikka?*  
where go-Q.DEF
- c. *eti ka-s(i).eyo?*  
where go-Q.POL
- d. *eti ka-yo?*  
where go-Q.POL
- e. *eti ka-si.o?*  
where go-Q.SEMF

5 Survey of the grammars of questions in Northeast Asia

f. *eti ka-o?*  
where go-Q.SEMF

g. *eti ka-si.na?*  
where go-Q.FAM

h. *eti ka-na?*  
where go-Q.FAM

i. *eti ka-si.e?*  
where go-Q.INT

j. *eti ka-(a)?*  
where go-(Q.INT)

k. *?eti ka-si.ni?*  
where go-Q.PLAIN

l. *eti ka-ni?*  
where go-Q.PLAIN

‘Where are you going?’ (L. Brown 2011: 47; Iksop & Ramsey 2000: 264f.;  
Song 2005: 158; Yeon & Brown 2011: 8)

Sohn (1999) includes the following Jeju example that corresponds functionally to the deferential speech level in the standard Korean example (176a) above.

(177) Jeju  
*etu ley ka-m-swu-kkwa?*  
where to go-IND-AH-Q.DEF

‘Where are you going?’ (Sohn 1999: 75, from Lee I.S.)

Saltzman (2014: 49) reanalyzed the sentence and calls *ley* (-*re* according to her) an ablative and *-swu* (-*su* in her rendering) a formal present tense marker, both of which are problematic. If *ley* indeed functions as an ablative, the sentence should rather have been translated as something like ‘Where do you come from?’ In fact, according to Sohn (1999: 75), Standard Korean may add the marker *lo* instead of *ley*. Clearly, this is the instrumental or directional case marker (*u*)*lo* and not an ablative (Song 2005: 115). A comparable sentence from Jeju in the past tense given by Kiaer (2014) is the following:

(178) Jeju  
*ödi ka-ng wa-m-su-gwa?*  
where go-PST ?AUX-IND-AH-Q.DEF

‘Where did you go?’ (Kiaer 2014: 10, my tentative analysis)

Here the marker *-m-su-gwa* is the same as *-m-su-kkwa* in (177) above and corresponds to the standard Korean deferential interrogative *-(su)p-ni-kka*. Note that *-sup* (-*p* when following a vowel) is an addressee honorific suffix, *-ni* is an indicative marker and only *-kka* is the actual question marker (Sohn 1994: 341). Thus, phonological differences apart,

Jeju *-m-swu-kkwa* (*-m-su-gwa*) and standard Korean *-sup-ni-kka* contain the same functional elements but apparently use the addressee honorific suffix and the indicative marker in reversed order.

Apart from Jeju, other dialects have special sentence enders as well. Table 5.63 summarizes those dialectal interrogative sentence enders that deviate from the standard language. Question marking in the Chungcheong dialect is very similar to Standard Korean, but *-o*, *-e-yo*, and *-sup-ni-kka* have the forms *-wu*, *-e-yu*, and *-sup-ni-kkya* instead, which exhibit slight phonological differences. Other endings such as *-nya* are identical:

- (179) Korean (Chungcheong)  
*ni pap mek-ess-nya?*  
 2SG meal eat-PST-Q.PLAIN  
 ‘Did you have your meal?’ (Sohn 1999: 71)

Table 5.63: Selected interrogative sentence enders in Korean dialects based on Sohn (1999: 66-76); some dialectal forms identical to standard forms were excluded; see also Yeon (2012)

	PLAIN	INT, FAM, SEMF	POL, DEF
Jeju	-(e)m-ti(ya), -(e)m-sini, -esinya	-m-kka, -m-kko, [-em-se], -(e)m-singa	-(e)m-swu-kkwa, (-wu)-kkwa
Hamgyong	-wa, -m	-wu, [-m-mey], [-cipi], -nungka, -m-twu(ng)	-sswu-ta, -m-mengi, -p-syo, -p-m-mi-kka
Pyongan	-(u)m-ma, [-(u)wa] <sup>20</sup>	[-(u)m-mey], -wu, -kan	-(u)op-ni-kka, -(su)p-ney-kka, -(su)p-mey-kka
Jeolla	-eya, -nya	[-elawu]	-(su)p-ni-kkye
Gyeongsang	PQ -na, COP -ka CQ -no, COP -ko	-neng-kyo	(-(si)p)-ni-kk(y)e
Standard	-ni, -(n)u-nya	[-e], -na, [-(u)o]	[-e.yo], -(su)p-ni-kka

Square brackets in Table 5.63 indicate forms that are not restricted to questions. Some examples from the dialects follow.

- (180) Korean (Hamgyong)  
*ka-wu?*  
 go-Q.FAM  
 ‘Does (she) go?’ (Sohn 1999: 67)

<sup>20</sup>The declarative form is *-(u)wa-yo*.

- (181) Korean (Pyongan)  
*etu.m ey ka-si-p-ney-kka?*  
 where to go-SH-AH-IND-Q  
 ‘Where are you going?’ (Sohn 1999: 69)

- (182) Korean (Jeolla)  
*ni pap muk-ess-nya?*  
 2SG meal eat-PST-Q.PLAIN  
 ‘Did you have your meal?’ (Sohn 1999: 74)

Several examples from the 19th century, mostly based on the Pyongan dialect (King 1987: 238), can be found in the *Corean Primer* by Ross (1877). For example, the ender *-um-mê* in (183) corresponds to *-(u)m-mey* in modern Hamgyong and Pyongan dialects.

- (183) Korean (Pyongan)  
*moosoon bäpi iss-um.mê?*  
 what meal COP-Q.FAM  
 ‘What food is there?’ (Ross 1877: 13)

For other dialects equally old materials are not available to me.

There are differences in intonation as well. In Jeolla and Chungcheong both falling and rising intonation are possible, whereas the standard Korean equivalent necessarily has rising intonation. Polar questions in Gyeongsang generally have a falling intonation. See Sohn (1999: 66-76) and Jeon (2015) for additional information.

Table 5.63 does not list forms encountered in **Yukcin** or **Kolyemal**. But some information on these dialects has been collected by Ross J. King. Instead of the standard Korean *-(su)p-ni-kka*, Yukcin has *-mdung* (King 1987: 238), which appears to have a cognate in Kolyemal *-(i)mdo ~ -mdu* (King 1987: 262). To my knowledge, no other Korean dialect mentioned thus far has a comparable form (Table 5.63). Kolyemal furthermore has *-na*, *-o*, and *-ja*, which correspond to Standard Korean *-na*, *-o*, and *-nya*, respectively. There are two polite markers, *-ga* and *-ge* that exhibit the same vowel difference as *-a ~ -e* in Standard Korean. But their exact etymology and function remain unclear to me.

- (184) Korean (Kolyemal)
- a. *misi-ř ha-ja?*  
 what-ACC do-Q.PLAIN  
 ‘What are you doing?’
  - b. *ka-mdo?*  
 go-Q.DEF  
 ‘Are you going?’
  - c. *ǰdi-ř ka-n.ga?*  
 where-ACC go-Q.POL  
 ‘Where are you going?’ (King 1987: 243, 262)

There is insufficient information on tag, focus, and alternative questions from the dialects. But like Standard Korean, almost all dialects have the same marking in polar and content questions. *Gyeongsang* is exceptional among modern dialects in making a distinction between polar *-no* and content questions *-na*. After copulas these markers take the forms *-ko* and *-ka*, but preserve the distinction between polar and content questions. This distinction cannot be found in more honorific speech levels.

(185) Korean (Gyeongsang)

- a. *ni etey ka-ss-no?*  
2SG where go-PST-Q.PLAIN  
'Where did you go?'
- b. *i ke nwu chayk i-ko?*  
this ?NOM who book COP-Q.PLAIN  
'Whose book is this?'
- c. *pap mun-na?*  
meal eat-Q.PLAIN  
'Did you eat?'
- d. *kuk i ni chayk i-ka?*  
that ?NOM 2SG book COP-Q.PLAIN  
'Is that your book?' (Sohn 1999: 72)

This pattern is a relic from Middle Korean that was lost in the other dialects during the **Pre-Modern Korean** period (Table 5.64). More exactly, the Middle Korean marker *-ko* was replaced by *-ka*, which from then on marked both polar and content questions (Sohn 2015: 456).

Table 5.64: Selected Pre-Modern Korean verb endings in the 19<sup>th</sup> century (Sohn 2015: 456)

	Statements	Questions	Commands	Proposals
Plain	-ta/-la, -eta, -ma	-nja/-njo, -lja	-la, -ala/ela	-ca
Intimate	-ci		-a	
Familiar	-ney, -lsjej, -msjej	-nka/-nko	-kej	-sej
Semi-formal	-(s)o	-o, -lka	-o, -kwulje	
Polite	-(j)o			
Deferential	-ita, -olsita, -oyta	-iska, -pnójka/-pnika	-sjosje, -psio	-saita, -psita

Several of these sentence enders still encountered in 19th century Korean are no longer in use in modern Standard Korean, e.g. the semi-formal interrogative ending *-lka*.

The difference between polar and content questions was still present in Middle Korean, which also had a further question marker *-ta* that was later lost. A good description of

Middle Korean question marking and its relation to Contemporary Korean (CK) was recently given by Sohn (2015: 448).

The interrogative endings were (a) *-(k)o/-sko*, (b) *-(k)a/-ska*, and (c) *-ta*. *-(s)ko* occur[red] in question-word question sentences, and *-(s)ka* in yes-no questions. Both *-(k)o* and *-(k)a* also attached directly to a copula complement, as in *i-nón sang-ka pel-a?* (CK *i-nun sang i-nka pel-inka?*) ‘Is this a prize or a punishment?’ After the mood suffixes *-ni* [indicative] and *-li* [prospective], the endings *-ko* and *-ka* lost the consonant *-k*, and became new question endings *-nio/-njo* and *-nia/-nja/-nje* on the one hand and *-lio/-ljo* and *-lia/-lja/-lje* on the other (CK *-n[j]a/-ni*; *-lya*). The question ender *-ta*, which is obsolete in CK, was frequently used in a sentence whose subject is a second person, as in *kutuj-nón enu cek-uj tolao-l-ta?* (CK *kutay-nun encey tolao-keyss-eyo?*) ‘When will you return?’ The three-way (a, b, c) distinction has been lost in CK, except that the Gyeonsang dialect retains the *-ko/-ka* distinction. (slightly corrected)

Consider the following examples that illustrate the markers *-ka*, *-ko*, and *-ta*, respectively.

(186) Middle Korean

- a. *i twu salóm i cinsillo nej hangkes-ka?*  
 this two person NOM truly 2SG.GEN master-Q  
 ‘Are these two persons truly your masters?’
- b. *hjenljang-ón sto mjes salóm-ko?*  
 wise.person-TOP also how.many person-Q  
 ‘Also, how many wise people were there?’
- c. *kutuj-nón enu cek-uj tolao-l-ta?*  
 2SG-TOP which time-LOC return-?PROS-Q  
 ‘When will you return?’ (Sohn 2012: 102, 103)

As can be seen from the example given in the above quotation, alternative questions take two polar question markers. For a better understanding, the example is analyzed in more detail in (187).

(187) Middle Korean

- i-nón sang-ka pel-a?*  
 this-TOP prize-Q punishment-Q  
 ‘Is this a prize or a punishment?’ (Sohn 2012: 102)

The complete set of Middle Korean sentence enders is given in Table 5.65. As in modern Korean, there are four different sentence types, but only four speech levels.

Old Korean had two interrogative sentence enders *-ku* 古, 遣, 故 and *-ka* 去, too, but both marked polar questions (Nam 2012: 58f.). The distinction between polar (*-ka* 去) and content (*-ko* 古, *-s.ko.a* 叱, 過) question markers was only introduced in Late Old Korean

(Nam 2012: 66). The Old Korean question markers display a form similar to Tungusic and Mongolic on the one side (Old Korean *-ku*) and to Japonic on the other (Old Korean *-ka*) (§§5.10.2, 5.8.2, 5.6.2). Question marking in the Jurchenic branch of Tungusic strongly differs from the other branches. There are more and different question markers and all have forms similar to Koreanic (Table 5.66).

Table 5.65: Middle Korean verb endings (Sohn 2015: 449)

	Statements	Questions	Commands	Proposals
Plain	<i>-ta/-la</i>	<i>-ko/-ka, -nje/-njo, -ta</i>	<i>-(ke)la</i>	<i>-cje (-cela)</i>
Neutral	<i>-ni/-noj</i>	<i>-ni</i>	<i>-kola/kolje</i>	?
Moderate respect	<i>-ng-ta</i>	<i>-nó-ni-ska/sko</i>	<i>-esje/-asje</i>	?
Deferential	<i>-ngi-ta</i>	<i>-nó/ni-ngi-ska/sko</i>	<i>-sjosje</i>	<i>-sa-ngi-ta</i>

Table 5.66: Similar question markers in Middle Korean and Jurchenic (§5.10.3)

Middle Korean	Manchu
<i>-ni</i>	=ni
<i>-nia/-nio</i>	=nio, ?Bala =η᠓
<i>-(k)a/-(k)o</i>	=o, Alchuka =(k)᠔
?-nja; ?Korean -na; ?Gyeongsang -nA	=nA

The exact source and time of borrowing remain unclear. But since Classical Manchu already had all markers, they were borrowed before 1600. A major difference is that question markers replace declarative endings in Koreanic but usually attach to them in Jurchenic (note, however, forms such as Bala *ənə=η᠓* ‘go=Q’). Manchu =o usually seems to follow copulas (free or bound), which also speaks in favor of a connection with Korean. Remember that the Gyeongsang dialect has the form *-ka* ~ *-ko* following copulas, and Alchuka =k᠔ preserves a velar plosive in this form as well. For instance, Manchu *-mbi=o* ‘-IPFV=Q’ (containing the copula *bi*) exactly corresponds to Alchuka *-mei=k᠔*. Similar to Korean sentence enders, Jurchenic markers may also attach to non-verbal elements but remain in sentence-final position, e.g. Bala *amin=η᠓* ‘father=Q’. Korean *-o* is not restricted to questions but may also mark imperatives, for instance, and Manchu also has a polite imperative marker *-rAo* that may contain the same element, possibly attached to the imperfective participle *-rA* that also appears in the prohibitive *ume V-rA*. But more research with the help of large scale corpora is necessary to determine the exact meaning and use of those markers in Manchu.

## 5.7.3 Interrogatives in Koreanic

Korean interrogatives exhibit two dominant resonances, *e~* and *m~*. The first has previously been compared with Old Japanese (§5.6.3). Similar to several other surrounding languages, the interrogative ‘who’ does not belong to any of these groups but rather starts with *n~*. In the Chungcheong dialect the resonance *e~* has the form *we~* instead (King 2006a: 267). Table 5.67 summarizes those interrogatives found in the literature available to me for Standard Korean, Korean as spoken in Jilin as well as Jeju.

Sohn (1999: 69) mentions a Pyongan interrogative verb *ekha* ‘to do how’ that he renders as the periphrastic sequence *etteh-key ha-* in standard Korean (*ha-* ‘to do’). Jeju has a periphrastic sequence  $\Delta\ddot{\Delta}\eta$ -*ha*, too (Saltzman 2014: 65). The interrogative meaning ‘who’ is an amalgamation of the original interrogative with the content question marker (Sohn 2015: 456). Note that *nwuku* still has the nominative form *nwu-ka* in Standard Korean. The combination *enu-cey* ‘which time’ is the source of the contracted form *encey* ‘when’ (Sohn 1999: 262).

Table 5.67: Interrogatives from Korean (Sohn 1999: 208ff., 256, 273, 396, 403; Yoon 2010: 2784, in square brackets), Korean spoken in China (Xuan Dewu et al. 1985: 29, 161), and Jeju (Kiaer 2014; Cheng & Harrison 2014, in square brackets; Saltzman 2014, in parentheses)

	South Korea	Jilin	Jeju
who	nwu(ku), nwu-ka NOM	nuku	[nuge]
what	mues	muyəʂ	musin’gǒ ~ musigǒ, [musinggeo]
what + noun	musun		
what + noun (time)	myech	mjətʃ	
which	enu	ənu	önügǒ, {əni}
what kind of	etten		
how	ecci, ettehkey		{ΔΔΔη}
where	eti	əti	ödi, {Δti}
when	encey	əntʃe	
how much	elma	ərma	
how long	[elmana]		
why	way		
what sort of	weyn		

There are also several forms meaning ‘who’ that are a combination of *enu-* with one of the three bound nouns *ay* ‘child’, *salam* ‘person’, and *pun* ‘respected person’ (Sohn 1999: 207f., passim; Song 2005: 73, passim). The interrogative *musun* is derived from *mues-i-n* ‘what-COP-REL’ and *etten* from *e-tte-ha-n* ‘which-kind-COP-REL’ (Sohn 1999: 256). Korean *weyn* similarly derives from *way-i-n*. A form without the relative marker *-n* but with the adverbializer *-key* ‘so that, to’ is probably the source of *etteh-key* ‘how’ (cf. Sohn 1999: 376). In these forms *e-* seems to be the actual interrogative marker that must also

be the ultimate source of *elma*, *eccī*, *enu*, and *eti*. In Jeju the interrogatives *musi(n')-gǔ* 'what' (Korean *mues*) and *ǫnǔ-gǔ* 'which' (Korean *enu*) contain a suffix *-gǔ* that could correspond to Korean *kes* 'thing' or *-kes-i* 'thing-NOM', which is regularly pronounced *-key* in informal speech (Song 2005: 155). This assumption is corroborated with data from *Kolyemal*, among which we find *misi-ge* 'what thing'. That in *Kolyemal* *ǫndze-ge* 'when' (Korean *encey*) the same suffix is present is unlikely from a functional perspective. Table 5.68 summarizes *Kolyemal* interrogatives and their direct Korean cognates.

Table 5.68: *Kolyemal* interrogatives in comparison with Korean (King 1987: 263; Sohn 1999)

	<b>Kolyemal</b>	<b>Korean</b>
who	nugi	nwu(ku)
who (ACC/DIR)	nugi-ř	nwukwu-l(ul)
how many	me(t)	myech
what (thing)	misi-ge	mues
what (ACC/DIR)	misi-ř	mues-ul (> mwe-l)
what kind of	musun	musun
where	ǫdi-mæ	eti
whither (ACC/DIR)	ǫdi-ř	eti-lo
whence (ABL)	ǫdi-sǫ	?eti-eyse
when (thing?)	ǫndze-ge	encey
why	ɔtʃtʃæ	eccī
how	ɔttɔk <sup>h</sup> æ	ettehkey

The resonance *e~* in Korean has the form *ɔ~* in *Kolyemal*. The case suffix *-ř* combines the function of a directive with that of an accusative, as can be seen in *nugi-ř* 'what-ACC' but *ǫdi-ř* 'where-DIR'. In Korean both the accusative *-(l)ul* and the instrumental *-(u)lo* also have the function of a directive, but the first is likely the source of *Kolyemal* *-ř* (Song 2005: 112, 115). *Kolyemal* *ǫdi-mæ*, like Pyongan *etu-m* in example (19) above, derives from Middle Korean *etu-mej* (see below).

Similar to Japanese (§5.6.3), Korean displays parallel paradigms in *demonstratives* and one interrogative stem. Like Japanese (*ko-*, *so-*, and *a-*, older *ka-*), Korean has a three way distinction of demonstratives (*i*, *ku*, and *ce*). But while Japanese has exactly the same paradigms for the interrogative stem *do-*, the paradigm of Korean *e-* exhibits several irregularities (Table 5.69).

The paradigms not only contain case endings but also certain bound nouns that have typological and probably areal parallels in Manchu (§5.10.3). Unlike the adverbs *yeki* 'here', *keki* 'there', and *yeki* 'over there', which are based on the demonstrative stems in combination with *eki* 'place', the interrogative *eti* 'where' has a case marker *-ti*. In Jeju, this suffix can also be found in the demonstratives (Table 5.70).

As regards the irregular Jeju stem *jo-*, note that Korean also has the diminutive demonstrative stems *yo*, *ko*, and *co* (Sohn 1994: 114).

5 Survey of the grammars of questions in Northeast Asia

Table 5.69: Full paradigms of Korean demonstratives and the selective interrogative (Sohn 1994: 296)

	PROX	PROX.H	DIST	which
stem	i	ku	ce	e-nu
<i>eki</i> ‘place’	<b>yeki</b>	<b>keki</b>	<b>ceki</b>	<b>eti</b>
<i>kes</i> ‘thing’	i kes	ku kes	ce kes	enu kes
<i>il</i> ‘thing, fact’	i il	ku il	ce il	enu il
<i>kos</i> ‘place’	i kos	ku kos	ce kos	enu kos
direction	i ccok	ku ccok	ce ccok	enu ccok
way	i-li	ku-li	ce-li	<b>eti-lo</b>
kind of	i-le-n	ku-le-n	ce-le-n	<b>e-tte-n</b>
to be ... way	i-leh-ta	ku-leh-ta	ce-leh-ta	<b>e-tteh-ta</b>
to do ... way	i-le-n-ta	ku-le-n-ta	ce-le-n-ta	-

Table 5.70: Jeju demonstratives and the selective interrogative in neutral and locative form (Saltzman 2014: 21)

	PROX	PROX.H	DIST	which
selective	i ~ jo/jΛ	ki	tçΛ	Λ-ni
locative	<b>jo-ti</b>	ki-ti	tçΛ-ti	Λ-ti

The interrogative *enu* ‘which’ is likely analyzable and based on the stem *e-*. The ending *-nu* might, according to Vovin (2005: 322), have a connection to a Japanese attributive ending (Old Japanese *-nō*). While the Jeju interrogative *Λti* ‘where’ can, at least synchronically, be analyzed as *Λ-ti* ‘which-LOC’, this is probably not true for Korean *eti*. Diachronically, however, both Jeju *Λti* and Korean *eti* go back to Middle Korean *e-tuj*, the second part of which is a bound noun meaning ‘place’. Vovin (2005: 322) assumes that the form can be reconstructed as Proto-Korean(ic) *\*əntúy*, thus allowing an analysis of the first part as the forerunner of Korean *enu* ‘which’ and a connection with Proto-Japonic *\*entu* ‘where’. His reasoning is based on the fact that the *t* should have regularly changed to *l* in this position without the *n* present. Middle Korean furthermore has an extended form *etu-mej* ‘where’ that might be comparable to Old Ryūkyūan *idu-ma* (§5.6.3).

In general, the set of Middle Korean interrogatives is very similar to modern Korean, only one form (*hjen* ‘how many’) having been entirely lost (Table 5.71). The exact differences between the forms meaning ‘what’ remain unclear to me.

Vovin (2005: 319) mentions an additional Middle Korean form *e:styé ~ e:sté ~ e:styéy* ‘how’ that, according to him, goes back to *\*e-is-ti* ‘how-exist-ADV’.

Table 5.71: Middle Korean interrogatives (Sohn 2015: 98) in comparison with modern Korean

	Middle Korean	Korean
who	nwu(-ko ~ -kwu)	nwuku ~ nwu-
how many	mjes	myech
what	musus, muzus, musu, musuk, musum	mues, musun
where	etuj	eti
where	etu-mej	etu-m (Pyongan)
whither	etu-le	eti-lo
which	enu, enwu, enó	enu
how much	enma	elma
when	enu-cjej	encey
how many	<b>hjen</b>	-

## 5.8 Mongolic

### 5.8.1 Classification of Mongolic

Mongolic languages form a language family with about a dozen modern members. According to Janhunen (2006: 232) they may be classified as in (5.4). Rybatzki (2003a: 388-389) assumes a slightly different classification with six groups. Of these, the so-called Northern (Khamnigan Mongol, Buryat) and South-Central groups (Shira Yughur) are part of Central Mongolic and Shirongolic, respectively, in Janhunen's (2006) classification.

The two classifications agree, however, in the number of languages as well as in some details such as the isolated positions of Dagur and Moghol. A recently discovered language that was added to (5.4) is called Kangjia and belongs to the Shirongolic branch. According to Kim (2003: 347), "the Kangjia 'language' would appear to be intermediate between Bonan and Santa." However, it may actually be more closely related to Bonan (Siqinchaoketu 2002: 66). Central Mongolic has also been called Common Mongolic by Janhunen (2012b: 3f.), and is said to contain also the Khorchin group of dialects that was not listed as a separate entry. Khorchin is spoken in western parts of Manchuria (the modern provinces of Heilongjiang, Jilin, and Liaoning), but mostly in the adjacent parts of eastern Inner Mongolia (Janhunen 2012b: 4). Chakhar Mongolian, not listed above, belongs to the same branch as Khalkha. It is spoken in Inner Mongolia and is said to be the language spoken by the descendants of the last emperor of the Mongolian Yuan dynasty and his followers who fled from Peking in 1368 (Sechenbaatar 2003: 1). Kalmyk, also not mentioned, can be considered an aberrant dialect of Oirat and is the only Mongolic language located in Europe. Moghol, located in Afghanistan, is probably extinct today and will for the most part be excluded here.

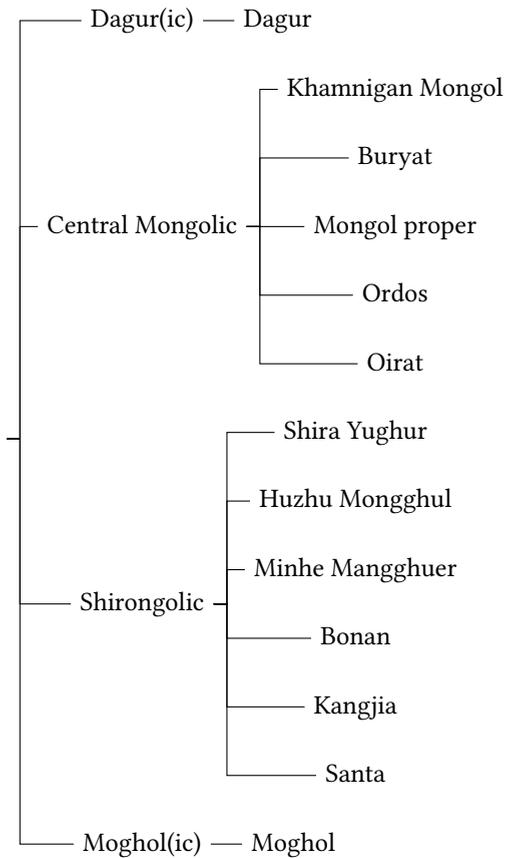


Figure 5.4: Classification of Mongolic

The Mongolic language Shira Yughur or Eastern Yughur (*dōngbù yùgù yǔ* 东部裕固语 in Chinese) should not be confused with the Turkic language Yellow Uyghur that is also called Sarig or Western Yughur (*xībù yùgù yǔ* 西部裕固语 in Chinese, see §5.11). There are also different Chinese designations for Bonan (*bǎoān yǔ* 保安语), Santa (*dōngxiāng yǔ* 东乡语), and a collective name for Huzhu Mongghul and Minhe Mangghuer (*tǔzú yǔ* 土族语), which are also known as Monguor in the West. Of the languages mentioned in (5.4) only Moghol is located outside Northeast Asia. All Mongolic languages except for Buryat and Kalmyk, which are for the most part spoken in Russia as well as Moghol in Afghanistan, are located within Mongolia and China.

### 5.8.2 Question marking in Mongolic

Question marking in Mongolic is not very complex. Janhunen (2003d: 27) gives a good summary of the marking of questions in Proto-Mongolic.

When no interrogative pronoun or pronominal verb was present in the sentence, *interrogation* in Proto-Mongolic was expressed by a sentence-final interrogative particle, which may be reconstructed as either \**gü* (> \*=*gU*), as in Buryat and Khamnigan Mongol, or \**xU* (> \*=*UU*), as in most other Mongolic languages. In questions containing an interrogative word, no particle was originally needed, but in Common Mongolic the copular form \**bü*-(*y*)*i* > \**büi* ‘being, present’ was grammaticalized in such sentences into what may be termed a *corrogrative* particle.

Consider the following examples from Written Mongolian in which the forms are still relatively well preserved.

(188) Written Mongolian

a. *ta sayin=uu?*  
2PL good=Q

‘Are you well?’

b. *ta ken bui?*  
2PL who COP>Q

‘Who are you?’ (Janhunen 2003f: 53, transcription changed)

The marking of polar questions with a sentence-final clitic is, of course, an areal trait. The development of a marker in content questions, on the other hand, sets Mongolic apart from most languages of the area. But similarly there exists a special content question marker in some Turkic languages that shares a functional background in a copula (§5.11.2 and §6).

In modern Mongolic languages the question markers have gone through phonetic erosion. As we will see further below the polar question marker fused with certain verb endings and copulas, especially in Shirongolic languages. Some individual Mongolic languages have additionally adopted question markers from other languages. Consider the following examples from a variety of Buryat spoken in China.

(189) Buryat (Shineken)

a. *xugs<sup>j</sup>in-tei oolz-aa=g=s<sup>j</sup>a?*  
old.woman-COM meet-P.IPFV=Q=2SG

‘Did you meet the old woman?’

b. *xen-tei oolz-aa=b=s<sup>j</sup>a?*  
who-COM meet-P.IPFV=Q=2SG

‘Whom did you meet?’

5 Survey of the grammars of questions in Northeast Asia

- c. *sʲii nam-tai ɔsʲ-nɔ=ba?*  
2SG 1SG.OBL-COM reach-PRS=Q

‘You go with me, don’t you?’ (Yamakoshi 2011a: 170-171)

Sentence (189a) illustrates the polar question marker =*go* ~ =*gu* ~ =*g*, sentence (189b) the optional “corrogrative” particle =*be* ~ =*b*, and sentence (189c) the marker =*ba*, which is a recent borrowing from Chinese *ba* 吧 that can be found in several languages of China (§6, and §5.9.2.1). In Shineken Buryat =*ba* is mutually exclusive with the agreement marker. Alternative questions display double marking with =*go* ~ =*gu* ~ =*g*.

- (190) Buryat (Shineken)

*bii enee-g-uur-ee jab-xa=g=bi, teree-g-uur-ee jab-xa=g=bi?*  
1SG.NOM this-E-INST-REFL go-P.FUT=Q=1SG that-E-INST-REFL go-P.FUT=Q=1SG

‘Should I go in this or in that direction?’ (Yamakoshi 2006: 153)

In non-verbal sentences the content question marker can also attach to word classes such as adjectives and interrogatives.

- (191) Buryat (Shineken)

a. *sʲinii xubuun=sʲe alin=be?*  
2SG.GEN boy.NOM=2SG.POSS which=Q

‘Where is your boy?’ (Yamakoshi 2011b: 116, shortened)

b. *alʻan=in hain=be?*  
which=3SG.POSS good=Q

‘Which one is good?’ (Yamakoshi 2007b: 5)

Janhunen (2003d) appears to believe that the question marker in Buryat and Khamnigan Mongol has a different origin than the one found in other Mongolic languages. Interestingly, both Buryat and Khamnigan Mongol had intense contact with dialects of the Tungusic language Evenki. In both Khamnigan Evenki and Khamnigan Mongol the enclitic has the form =*gv*. Janhunen (1991: 95) speculated that it may have been borrowed from one language to the other, but left the direction of borrowing open. Given that many Tungusic languages preserve a cognate of the enclitic in Khamnigan Evenki (see §5.10.2), it seems likely that it was borrowed from Evenki into Khamnigan Mongol. But Khamnigan Evenki may reflect influence from Khamnigan Mongol, and in turn has lost the property of consonant alternation that is still present in Evenki proper (= *gu* ~ = *ku* ~ = *ɲu* ~ = *vu*). The enclitic =*gi(i)* in the Tungusic language Solon, on the other hand, is probably a secondary loan from a Mongolic source (possibly Buryat =*giü*). Apart from the not unlikely scenario that individual Mongolic languages have borrowed the Tungusic question marker, the other Mongolic question marker reconstructed by Janhunen as \**xU*, could potentially also have a very old connection to Proto-Tungusic \**Ku* because it already existed at the proto-level of both language families. As is often the case, the etymology of the markers is not transparent in either Mongolic or Tungusic. Also note a similar marker -*ku* (written as 古, 遣, 故) in Old Korean (§5.7.2).

The form of the question marker in Middle Mongol was probably =UU, that is =üü ~ =uu. In Written Mongol (Uyghur script), the enclitic has the form -(ju)gu ~ -(ju)qu ~ -(ju)qhu when following vowels and -ugu ~ -uqu ~ -uqhu otherwise (Rybatzki 2003b: 79). According to Street (2008/09: 45), the plosives were not present in the spoken language but rather indicated a hiatus, which can be seen from other scripts used to write Middle Mongol. The vowel harmony may represent a problem for the comparison with Tungusic, but the older records of Middle Mongol show a strong functional similarity to Tungusic. While the enclitic has a strict sentence-final position in modern Mongolian, it was mobile at earlier stages and could attach to a focused element. In other words, the functional scope included not only polar but also focus questions.

The interrogative particle in early Middle Mongolian was what may be termed a **floating particle**: for purposes of emphasis it could float from one point to another on the surface structure of a sentence, though at a deeper level remaining in construction with the remainder of the sentence as a whole [i.e., marking the whole sentence as question]. (Street 2008/09: 76, my square brackets)

A typological parallel for a change from a mobile to a sentence-final question particle can be observed in the transition from Old to Modern Japanese (§5.6.2). In Middle Mongolian alternative questions were also marked with the same enclitic that attached once on each alternative. Consider the following examples from Middle Mongol.

(192) Middle Mongol (Arabic script; Secret History)

a. *burut-b=uu?*

escape-TERM=Q

‘Did (you) escape?’

b. *caq=u’u gür-be?*

time=Q arrive-TERM

‘Has *the time* arrived?’

c. *jöb=ü’ü          tab=u’u          ügüle-rün?*

appropriate=Q ?convenient=Q say-CVB

‘Saying: is it appropriate, is it convenient?’ (Rybatzki 2003b: 79)

The same functional scope can be reconstructed for Proto-Tungusic (§5.10.2). Furthermore, the two proto-languages combine this with a similar phonological shape, which is unlikely to be a coincidence.

As indicated by Janhunen in the above quotation, the etymology of the marker \**büi* is transparent and has its origin in a participle form of the copula \**bü-*, most likely the so-called deductive \*-(y)i ‘PRS.IPFV’ (Janhunen 2003d: 24). The term “corrogrative” is frequently employed by Janhunen but has never been explained adequately from a functional perspective or in terms of grammaticalization. According to the analysis followed in this book, it may simply be called a content question marker. While in Mongolian it has an eroded form similar to Buryat, it may also appear in a form that is still identical to the copula.

5 Survey of the grammars of questions in Northeast Asia

- (193) Mongolian  
*cii xedzee yab-sen bwai?*  
2SG when depart-PST COP>Q  
'When did you go?' (Janhunen 2012b: 255)

As noted above, a similar content question marker exists in some surrounding Turkic languages that has its origin in a copula that in turn goes back to a demonstrative (§5.11.2).

**Dagur** differs from other Mongolic languages in that there is a different polar question marker. There is no, or at least no obligatory, content question marker.

- (194) Dagur
- a. *en bitig=yee?*  
this book=Q  
'Is this a book?'
  - b. *fii ani-fi?*  
2SG who-2SG  
'Who are you?' (Tsumagari 2003: 150; Chaolu Wu 1994b: 11)

The data by Zhong Suchun (1982), collected in 1963 in Morin Daba, show a similar situation but make it clear that the polar question marker can be used optionally in content questions, too. This indicates that it is not only formally, but also functionally different from other Mongolic languages. Apparently, Dagur also has borrowed Chinese *ba* 吧.

- (195) Dagur (Morin Daba)
- a. *ənə warkəl-ii waa-səŋ=jəə?*  
this clothes-ACC wash-PST=Q  
'Have (you) washed these clothes?'
  - b. *xaanə itf-bəi-fii=jəə?*  
whither go-FUT-2SG=Q  
'Where are you going?'
  - c. *danŋ oo-dʒ ul bol-səŋ baə?*  
tobacco drink-CVB.IPFV NEG can-PST Q  
'Smoking is prohibited here, right?' (Zhong Suchun 1982: 76)

There is one example of an alternative question that exhibits the marker *jumoo* once on each alternative. This is probably a recent loan from an Inner Mongolian dialect, in which the latter part is the question marker =*UU*, that will be further explained below.

- (196) Dagur  
*bii əidəə jaw-oos-min<sup>i</sup> dʒuyi-yu jum.oo? tiidaa jaw-oos-min<sup>j</sup>*  
 1SG this.way go-CVB.COND-1SG right-IPFV Q that.way go-CVB.COND-1SG  
*dʒuyi-yu jum.oo?*  
 right-IPFV Q  
 ‘Will I go this way or that way?’ (Chaolu Wu 1994b: 18)

In the Dagur dialect spoken in Tarbagatai (*tǎchéng* 塔城) in Xinjiang, the usual polar question marker has the vowel-harmonic forms *-ja ~ -jə ~ -jo* and also marks alternative as well as content questions. It remains unclear whether it can also be found in focus questions. The marker was given as a suffix but is reanalyzed as an enclitic here.

- (197) Dagur (Tacheng)
- a. *ərgun-šin xajir-səŋ=ja?*  
 husband-2SG come.back-PST=Q  
 ‘Did your husband come back?’
  - b. *səəs-šin xərəə, bəraan=ja očog=jo?*  
 urine-2SG how.much much=Q little=Q  
 ‘How much urine do you have? Is it a lot or a little?’
  - c. *šii xəjəə ənd ir-səŋ-ši=jə?*  
 2SG when here come-PST-2SG=Q  
 ‘When did you come here?’ (Yu Wonsoo et al. 2008: 85, 86)

The functional scope of the question marker in Dagur suggests an areal connection to several surrounding languages (§6).

In some content questions there is a copula that could be the “corrogiative” form found in other Mongolic languages. As in Shineken Buryat the agreement marker follows the copula, but in Dagur the sentence additionally takes the usual question marker, which makes it unlikely that the copula fulfills the role of a question marker.

- (198) Dagur (Tacheng)  
*šii xaan-aar ir-səŋ-b-ši=jə?*  
 2SG where-ABL come-PST-COP-2SG=Q  
 ‘Where did you come from?’ (Yu Wonsoo et al. 2008: 86)

According to Yu Wonsoo et al. (2008: 79), the marker *-jə* sometimes fuses with the preceding suffix and the verb in example (198) is realized as */irzbišə/*. If the element *=jə* that is sometimes found in questions in the Tungusic languages Sibe and Aihui Manchu is indeed a question marker, then its most likely source is Dagur. Clearly, Dagur was also the origin of the question marker *=jee* in Oroqen (see §5.10.2).

To my knowledge there are no explicit descriptions of questions in **Moghul**, but Weiers (1972) mentions several examples of polar and content questions that appear to be generally unmarked morphosyntactically. Presumably, there was a different intonation con-

tour that cannot be reconstructed for now. Given its peripheral position outside of Northeast Asia, Moghol will not be further addressed here.

The Northern subgroup of Mongolic as identified by Rybatzki (2003a), i.e. Khamnigan Mongol and Buryat, basically share the question marking of Shineken Buryat seen above. Both the polar question marker as well as the “corrogrative” particle are still present in both languages. Khamnigan also has adopted the Mandarin marker *ba* 吧. The Khamnigan Mongol “corrogrative” particle *bei* has been borrowed into Khamnigan Evenki (§5.10.2).

(199) Khamnigan Mongol

- a. *hain=gu?*  
good=Q  
‘Is it good?’
- b. *tere ken bei?*  
3SG who Q  
‘Who is he?’ (Janhunen 2003b: 97)
- c. *ənə kobcaxon=cini tɔrg-ɔɔr ɔjɔ-gd-ɔɔ=gu, bisi=gu?*  
this clothes=2SG.POSS silk-INST sew-PASS-P.IPFV=Q NEG=Q  
‘Are these clothes made of silk or not?’
- d. *bii kara-ku-du=min tabi-tee-ta ba?*  
1SG.NOM see-P.FUT-DAT=1SG.POSS fifty-PROP-2PL Q  
‘I guess you are about fifty years old, right?’ (Yamakoshi 2007a: 132, 127)

As in Shineken Buryat, an agreement suffix may follow the question markers in Standard Buryat. The marker =*gü* marks polar, alternative, and maybe focus questions.

(200) Buryat

- a. *yeshe münge-tei hen=gü?*  
PN money-COM COP.PST=Q  
‘Did Yeshe have money?’
- b. *shi xen-tei=b=shi?*  
2SG who-COM=Q=2SG  
‘Who are you with?’
- c. *yaba-xa-m=gü, bai-xa-m=gü?*  
go-P.FUT-1SG=Q stay-P.FUT-1SG=Q  
‘Shall I go or stay?’ (Skribnik 2003: 120, 119)

This is probably also true for Khamnigan Mongol, but no example for a plain alternative question has been found in the relevant literature.

In order to compensate for the lack of information in most grammatical descriptions, the following examples of Cyrillic Khalkha Mongolian were elicited in October 2015

from a Mongolian informant of Outer Mongolia living in Germany. The analysis and transcription partly follows Janhunen (2012b). As noted before, polar questions are usually marked with the enclitic =UU.

- (201) Cyrillic Khalkha Mongolian  
*ci surguul-ruu-g.aa yaw-j bai-g.aa youm=uu?*  
 2SG school-DIR-POSS.REFL depart-CVB.IPFV COP-P.IPFV COP=Q  
 ‘Are you going to school?’

As in this example (201), the enclitic sometimes combines with a copula, derived from a word meaning ‘thing’ (Janhunen 2012b: 221, 228). This form also appears in Dagur as =*jumoo* and some Tungusic languages (see §5.10.2), all of which were probably borrowed from central Mongolian dialects spoken in Inner Mongolia. It also seems likely that the polar question marker found its way from Mongolian (=uu ~ =oo) into Oroqen (=oo), where it has an additional meaning of fear or doubt. Focus questions are identical to polar questions in form but exhibit an additional intonational peak on the focused element (indicated by underlining in Mongolian and with italics in the translation). Unlike Middle Mongol and some Tungusic languages, the question marker does not express focus itself and cannot take any other position in the sentence.

- (202) Cyrillic Khalkha Mongolian  
*ci surguul-ruu-g.aa yaw-j bai-g.aa youm=uu?*  
 2SG school-DIR-POSS.REFL depart-CVB.IPFV COP-P.IPFV COP=Q  
 ‘Are you going to school?’

Both plain and negative alternative questions require two question markers as well as a disjunctive. The disjunctive *eswel* literally meaning ‘(and) if not’ could be analyzed as *es-wel* ‘NEG-CVB.COND’ and can also be employed as a standard disjunctive (Janhunen 2012b: 221). This has a typological parallel in Korean *an-i-myen* ‘NEG-COP-COND’ (§5.7.2).

- (203) Cyrillic Khalkha Mongolian
- a. *ci tzai uu-x=uu, eswel airag uu-x=uu?*  
 2SG tea drink-P.FUT=Q or kumis drink-P.FUT=Q  
 ‘Do you drink tea or kumis?’
- b. *ci surguul-ruu-g.aa yaw-a.x=uu eswel yaw-a.x-güi=y.uu?*  
 2SG school-DIR-POSS.REFL depart-P.FUT=Q or depart-P.FUT-NEG=Q  
 ‘Do you go to school or not?’

Alternative questions may also take the extended question marker *youm=uu* (Benjamin Brosig p.c. 2016).

- (204) Cyrillic Khalkha Mongolian  
*ci tzai uu-x youm=uu, eswel airag uu-x youm=uu?*  
 2SG tea drink-P.FUT COP=Q or kumis drink-P.FUT COP=Q  
 ‘Do you drink tea or kumis?’

5 Survey of the grammars of questions in Northeast Asia

Apparently, the question marker =UU has expanded its scope and sometimes also appears in content questions.

- (205) Cyrillic Khalkha Mongolian  
*ci xedzee surguul-ruu-g.aa yaw-a.x=uu?*  
 2SG when school-DIR-POSS.REFL depart-P.FUT=Q  
 ‘When are you going to school?’

But according to other sources, Khalkha also has the expected “corrogrative” particle.

- (206) Cyrillic Khalkha Mongolian  
*xen tsai uu-san be?*  
 who tea drink-P.PFV Q  
 ‘Who drank tea?’ (Svantesson 2003: 171)

Some verbal endings in Mongolian have a slightly different but predictable form in the interrogative than those in the declarative. These are summarized in Table 5.72.

Table 5.72: Special interrogative endings in Mongolian according to Janhunen (2012b: 183f., 255, 298); differences are marked with boldface

	Meaning	Declarative	Interrogative
Mood	VOL	yab- <b>ii.y</b>	yab-y=oo
Finite	DUR	yab- <b>e.n</b> ’	yab- <b>n</b> =oo
	CONF	yab-l(=aa) EMPH	yab-l=oo/yab-laa=y.oo
	TERM	yab- <b>eb</b>	yab- <b>b</b> =oo
	RES	yab-j	yab-j=oo
Participle	FUT	yab- <b>ex</b>	yab- <b>x</b> =oo
	HAB	yab- <b>deg</b>	yab- <b>dg</b> =oo
	PRF	yab- <b>sen</b>	yab- <b>sn</b> =oo
	IMPRF	yab-aa	yab-aa=y.oo

Similarly to other languages of the region, descriptions of Mongolic languages usually do not mention tag questions and it remains open whether they are absent or were simply ignored. The following elicited example is marked with a marker *tee* that appears to be ultimately derived from the distal demonstrative *te-* and can roughly be translated as ‘is it like this?’.

- (207) Cyrillic Khalkha Mongolian  
*ci surguul ruu-g.aa yab-na tee?*  
 2SG school DIR-POSS.REFL depart-DUR SO  
 ‘You are going to school, right?’

Another tag question type encountered in Mongolian consists of a negative copula followed by a polar question marker.

- (208) Darkhat Mongolian  
*ir-sen        biš=oo?*  
 come-P.PFV NEG=Q  
 ‘(S)he arrived, didn’t (s)he?’ (Ragagnin 2011: 188)

Descriptions of Mongolic languages usually also do not mention intonation contours. But Karlsson (2003: 192) made the following interesting observations for Khalkha Mongolian.

Focus in questions is signaled by a rising gesture, the LH [low high]. However, depending on the segmental conditions, the gesture can be realized just as a tonal peak, synchronized with the second mora, making it similar to the focal H in declaratives. Interrogatives have a terminal low boundary tone, which is characteristic for most informants, while the high final rise is optional. All this makes the intonation of interrogatives similar to that of declaratives. The reason for this seems to be the strong formal signaling of interrogatives by using question particles. Thus, intonation has a redundant role in forming the interrogative mode in Mongolian. (my square brackets)

We may thus conclude the following: polar questions are obligatorily marked with the enclitic =*UU* and have an optional rising intonation. In focus questions there is an additional peak on the focused element. This makes the structure of focus questions quite different from Middle Mongol, where, as seen above, the enclitic attaches to the element in focus. In addition, interrogatives in content questions obligatorily receive “the same tonal gesture” as focused elements in focus questions (Svantesson et al. 2005: 93).

In **Chakhar Mongolian** the polar question marker is =*UU* (= *ũũ*, =*uu*) or =*γ.UU* when following a vowel. According to Sechenbaatar (2003: 182) “the material shape of the interrogative particle links Chakhar with Khalkha, marking a distinction with regard to several other Inner Mongolian dialects, such as, for instance, Baarin, in which the interrogative particle appears in the invariable shape =*ii*.” The optional “corrogrative” particle has the form =*w* ~ =*b* or =*wéé* ~ =*béé*. The forms with a plosive are found following the nasals *m* or *ŋ*.

- (209) Chakhar Mongolian
- a. *ax=cin        õnõödõr yaw-n=uu?*  
 brother=2SG today depart-DUR=Q  
 ‘Is your brother leaving today?’
  - b. *e.n yamar yum=béé?*  
 this what thing=Q  
 ‘What kind of thing is this?’ (Sechenbaatar 2003: 182, 107)

**Khorchin Mongolian** likewise has the enclitic  $=\langle j \rangle UU$  that marks polar, alternative, and possibly focus questions.

(210) Khorchin Mongolian

a. *ənə bəl nəm=ʋʋ?*

this TOP book=Q

‘Is this a book?’

b. *bii ənuur jab-aal taar-n=ʋʋ, tiiš-əə jab-aal taar-n=ʋʋ?*

1SG here go-CVB.COND fit-prs=Q there-REFL go-CVB.COND fit-PRS=Q

‘Do I have to go in this or that direction?’ (Yamakoshi 2015: 282, 292)

There is also an enclitic  $=\langle j \rangle ii$  that marks polar questions as well as, optionally, content questions. This might indicate an areal connection to Ainu, Dagur, Korean, Japanese, Manchu, Ōgami, and Ulcha (§6). Perhaps, the expansion of Khalkha  $=\langle y \rangle UU$  can also be explained as an areal trait connected to this.

(211) Khorchin Mongolian

a. *ter xun jeb-lɛɛ=j.ii?*

that person depart-CONF=Q

‘Did that man go?’ (Chaghanhada 1991: 71)

b. *čii xən=ji?*

2PL who=Q

‘Who are you?’

c. *činii aab=čin’ xaa bii?*

2SG.GEN father=2SG.POSS where COP

‘Where is your father?’ (Yamakoshi 2015: 282, 284)

However, Khorchin might also exhibit the “corrogrative” marker. Compare the following two examples from Khorchin and Khalkha, respectively (Benjamin Brosig p.c. 2018).

(212) Khorchin Mongolian

*ən tɛxaa xən-ɛɛ jimɛɛ?*

this chicken who-GEN COP.Q

‘Whose chicken is this?’ (Chaghanhada 1991: 71)

(213) Cyrillic Khalkha Mongolian

*en taxyaa xən-ii(-x) youm bwai?*

this chicken who-GEN(-NOM) COP Q

‘Whose chicken is this?’

Without the nominalizer, *youm* is perhaps better understood as ‘thing’.

According to Brosig (2014: 15), Khorchin has two further question markers =*me* and =*mu*. Their exact scope and etymological relation remain unclear. However, =*me* apparently can mark polar and content questions while =*mu* appears at least in polar and alternative questions, e.g. *nogon=mu, xar=mu*? ‘Is (it) green or black?’ (Brosig 2014: 15).

- (214) Khorchin Mongolian  
*zaqi-d yuu xii=me?*  
 PN-DAT what do=Q  
 ‘What are you going to do in Jarud?’<sup>21</sup> (Brosig 2014: 15)

An imperfective marker *-n* is said to have been assimilated to the following question marker in this example. Khorchin also has borrowed the Mandarin marker *ba* 吧. According to Chaghanhada (1991: 72) it has a long vowel (*baa*), just like the adjacent languages Solon, Oroqen, and Dagur.

- (215) Khorchin Mongolian  
*činii ax=čin’ bas duč bəl-ədue=ba?*  
 2SG.GEN e.brother=2SG.POSS also forty become-NEG.IPFV=Q  
 ‘Your elder brother is not yet forty, right?’ (Yamakoshi 2015: 287)

An autochthonous equivalent of Mandarin *ba* 吧 used mostly by older speakers is the combination =*i=dee* (Brosig 2014: 16). In tag questions either *ba* or =(y)UU may be employed, which has parallels in the Tungusic language Sibe (§5.10.2) and in Mandarin (§5.9.2).

- (216) Khorchin Mongolian  
 a. *tərə xun baxš bišə, tiim=ba?*  
 that person teacher NEG like.that=Q  
 ‘That person is not a teacher, right?’  
 b. *xədən čag bəl-ǰ=ji? xəjər čag=ʊʊ? tiim=uu?*  
 when hour become-CVB.IPFV=Q two hour=Q like.that=Q  
 ‘What time is it? Two o’clock? Right?’ (Yamakoshi 2015: 283, 286)

Benjamin Brosig (p.c. 2016) mentions a couple of additional particles such as *qi* (identical to *ji* below) of not absolutely clear origin. It is perhaps best classified as a tag question marker. Mongolian has a negative copula, *bish* in the spoken and *bous* in the literary language, that might somehow be connected to a word meaning ‘other’ (Janhunen 2003d: 27; Janhunen 2012b: 251). There is a parallel grammaticalization of adjectives meaning ‘different’ to a negative copula in Tungusic (Hölzl 2015a: 146). According to Chaghanhada (1991) it has the form *bifii* in Khorchin and has developed into a question marker. Under my analysis, however, the final =*ii* might be nothing but the question marker. From this perspective, *bif=ii* is probably a tag question marker almost identical to Darkhat *biš-oo*

<sup>21</sup>In Chinese this place is called *zāqi* 扎旗.

5 Survey of the grammars of questions in Northeast Asia

and *bish=uu* in Mongolian according to Janhunen (2012b: 251). This construction has exact typological parallels in several Tungusic (§5.10.2) and Turkic languages (§5.11.2). In what appears to be another type of tag question, Khorchin *bif* may also be followed by an emphatic enclitic (*bif=j.əə*) that has the form *=(y)AA* in Mongolian according to Janhunen (2012b: 93). Perhaps the form *fii* is a contracted form of *bif=ii*.

- (217) Khorchin Mongolian  
*ən udur fin tabən sar-iin nəgən fii?*  
 this day new five month-GEN one Q  
 ‘Is today not the first day of May?’ (Chaghanhada 1991: 71)

The same marker *fii* can also be found in tag questions following the element *tii.n*, which is probably derived from the distal demonstrative (cf. Janhunen 2012b: 130), similar to (207) from Khalka. Chaghanhada (1991: 72) translates *tiin fii*, which may be attached to a declarative sentence, as a tag question. In Khalkha there is also a question tag *tiim bish=üü* (Benjamin Brosig p.c. 2018).

There are few clear descriptions for questions in **Ordos**. But there is evidence that it preserves the original question marker as *=(j)uu* and lacks the “corrogrative” particle (Stefan Georg p.c. 2015).

- (218) Ordos  
 a. *yabu-b=uu?*  
 go-TERM=Q  
 ‘Did he go?’ (Georg 2003b: 208)  
 b. *t‘e.re juuu ge-Džē-n?*  
 3SG what say-RES-3  
 ‘What does he say?’ (Mostaert 1937: lix)

Most other dialects will be ignored here for lack of data and reasons of space.

There are also few good materials for questions in Oirat, which is why there will first be a descriptions of questions in the closely related language (or aberrant dialect) Kalmyk. In *Kalmyk* the interrogative particle *=u* marks polar questions and similar to Mongolian (Table 5.72) fuses with some suffixes, e.g. *-na* ‘DUR’ vs. *-nu* ‘DUR.Q’ and *-la* ‘CONF’ vs. *-lu* ‘CONF.Q’ (Benzing 1985: 42). The “corrogrative” particle is preserved as *=b ~ =w*, e.g., *kem=b?* ‘who is it?’. In addition, there is another question marker *=iy ~ =i* that seems to be employed in alternative questions as well as polar questions, e.g. *xol=iy?* ‘Is it far?’.

- (219) Kalmyk  
 a. *ter ir-v=u?*  
 3SG come-PFV=Q  
 ‘Did he come?’  
 b. *endr yamaran ödr? sän ödr=iy mu ödr=iy?*  
 today which day good day=Q bad day=Q  
 ‘What kind of day is it today, (is it) a good day or a bad day?’ (Benzing 1985: 42f.)

Note that in this example a content question is followed by an alternative question (see §4.4). Bläsing (2003) does not mention the marker, but quite clearly, this is the same element we have already seen above, e.g. Khorchin  $=(\textit{j})\textit{ii}$ . The following Kalmyk sentences were elicited from a native speaker living in Germany in January 2016 via internet. The transliteration and analysis are mine but roughly follow Bläsing (2003).

(220) Kalmyk

- a. *al'daran* *yow-jana-c?*  
whither go-PROG-2SG  
'Where are you going?'
- b. *ci manhdur shköl-d yow-jana-c?*  
2SG tomorrow school-ALL go-PROG-2SG  
'Are you going to school tomorrow?'
- c. *ci manhdur shköl-d yow-jana-c?*  
2SG tomorrow school-ALL go-PROG-2SG  
'Are you going to school tomorrow?'

No question marker appears in content questions. In focus questions the focus is apparently expressed with the help of an additional peak on the focused element.

The situation in Kalmyk is indeed very similar to Oirat proper, for which Birtalan (2003) mentions the question markers  $=UU \sim =(y)UU$ ,  $=ii$ , as well as  $=w \sim =b$ .

(221) Oirat

- sään bään=uu?*  
good COP=Q  
'Are you well?' (Birtalan 2003: 227)

As opposed to Benzing, she treats the form  $=ii$ , which we have already encountered in Baarin, Khorchin, and Kalmyk, as a variant of the polar question marker. In fact, this is the most likely analysis as its form  $=ii \sim =y.ii$  is completely parallel to the standard marker  $=UU \sim =y.UU$  (Janhunen 2012b: 183). In some Tungusic languages there are question markers that were probably borrowed from Mongolian  $=(y)ii$ , notably Ongkor Solon  $-ii$  as well as, less likely due to geographical distance, Even  $-ii$ , Negidal  $-i$ , and maybe Uilta  $-(y)i$  (§5.10.2).

Shirongolic languages also preserve the original polar question marker, but display a more complicated picture than Central Mongolic. In *Shira Yughur*—classified by Rybatzki (2003a) as the only South-Central language instead of as Shirongolic—the polar question marker  $=uu \sim =j.uu$  sometimes fused with the preceding verb ending, but there is no clear information as to when and how often this happened. The durative marker  $-nAi$  (and variants) always has the form  $-nam$  before the question particle. The “corrogate” particle appears to be optional. Alternative as well as negative alternative questions take two question markers.

5 Survey of the grammars of questions in Northeast Asia

(222) Shira Yughur

a. *tʃə qutad largə ʃda-dac=uu?*  
2SG PN speak able-HAB=Q

‘Can you speak Chinese?’

b. *tʃəmiin dʒa jimar dʒe-də ɔl-sən bə?*  
2SG.GEN be what time-DAT get-PST Q

‘Which year were you born in?’ (Chaolu Wu 1994d: 8, 9f.)

c. *tʃə.ma-də qudaxa bii=j.u, uyui=j.u?*  
2SG.OBL-LOC knife EX=Q NEG=Q

‘Do you have a knife?’ (Zhaonasitu 1981a: 58)

The last example (222c) is a negative alternative question that shows a negative existential because of the existential in the first alternative. Similar to the situation in Khalkha before, there is also one example of the polar question marker in what appears to be a content question (cf. Mongolian *-x=oo* in Table 5.72 above).

(223) Shira Yughur

*cimiin keen-di ög-k’uu?*  
2SG.ACC who-DAT give-FUT.Q

‘To whom shall I give you?’ (Nugteren 2003: 280)

In sum, Shira Yughur interrogative constructions pattern with Central Mongolic and have to be differentiated from the more complex system found in Shirongolic languages.

*Bonan*, like *Ordos*, lacks the “corrogrative” particle in content questions, which are morphosyntactically unmarked.

(224) *Bonan*

*dzoma xala o-to?*  
PN where go-PFV

‘Where did Droma go?’ (Fried 2010: 261)

For polar questions *Bonan* preserves the Mongolic interrogative marker that has the form *-u*. But its use is more complicated than in these Mongolic languages we have encountered before: “When *-u* is suffixed to imperfective lexical verbs, it replaces the imperfective suffix (*-tci/-tco*). Similarly, when it is suffixed to perfective verbs, it replaces the perfective suffixes *-to* and *-tə*.” (Fried 2010: 258) The suffix thus attaches directly to the verb stem.

(225) *Bonan*

*tʃʰə nudə natʰə-u?*  
2SG today dance-Q

‘Did you dance today?’ (Fried 2010: 259)

Table 5.73: Special interrogative forms in Bonan (Hugjiltu 2003: 339, 343)

	Meaning	Declarative	Interrogative
NARR	PRS, FUT	-m	-mu
DUR	PRS, FUT	-na	-nu
TERM	PST	-wa > -o	-wu > -u

Table 5.74: Special interrogative copula forms in Bonan (Hugjiltu 2003: 340, 343)

	Subjective	Objective	Interrogative
EXIST	wi	wa	wu
COP.EMPH	mbi	mba	mbu

The interrogative marker *-u* fused with several verb endings and copulas, see Table 5.73 and Table 5.74.

(226) Bonan

- a. *ode-nu?*  
go-DUR.Q  
'Will he go?' (Hugjiltu 2003: 343)
- b. *tcin-da tca dawzi wəu?*  
2SG-DAT tea leaf EXIST.Q  
'Do you have tea?'
- c. *tə.rə aagə da təma mbəu?*  
that girl also student COP.EMPH.Q  
'Is that woman also a student?' (Chaolu Wu 1994a: 8, 6)

The copula forms are given as *wəu* and *mbəu* by Chaolu Wu (1994a). According to Fried (2010: 260), the forms are declarative *wi* SUBJ vs. *wa* OBJ, interrogative *wu(u)* SUBJ vs. *wa-u* OBJ and declarative *bi* SUBJ vs. *ba* OBJ, interrogative *bu* SUBJ vs. *ba-u* OBJ. The copula starting with *b-* is used in nominal copula clauses, the copula starting with *w-* in all other clause types (Fried 2010: 260). In addition, the Gansu variety of Bonan has borrowed the Chinese polar question marker *ma* 吗 (Hugjiltu 2003: 343) and has a special marker *-si*, allegedly for “rhetorical questions”, that can also mark alternative questions.

(227) Bonan

- a. *χecəŋ jaŋgətə nat<sup>h</sup>ə k<sup>h</sup>ər-si?*  
PN how dance be.required-Q  
'How should (one) dance Leru?'

5 Survey of the grammars of questions in Northeast Asia

- b. *pə [hkuətə orə-si əsə orə-si] əsə med-o.*  
 1SG yesterday rain-Q NEG rain-Q NEG know-TERM  
 ‘I don’t know [whether it rained or not yesterday].’ (Fried 2010: 99, 227)

An example of a plain alternative question was given by Chaolu Wu (228). Interestingly, only the first alternative has an overt question marker. A similar situation can be seen in Santa, Kangjia, and Mangghuer and is an areal feature.

- (228) Bonan  
*bə en-sa ə.d kə-saŋ=wu, taŋ-sa ə.d kar-saŋ?*  
 1SG this-ABL go be.required-P.PFV=Q that-ABL go be.required-P.PFV  
 ‘Will I go this way or that way?’ (Chaolu Wu 1994a: 15)

In Bonan there is also the Chinese question marker *ba* 吧.

- (229) Bonan  
*dedə ‘gudə sə edəro ba?*  
 old.man.VOC yesterday NEG ?tired Q  
 ‘Grandpa, you weren’t too tired yesterday, right?’ (Buhe & Liuzhaoxiong 1982: 59)

In **Kangjia** the question marker has the form *-u* and has the two variants *-vu* and *bu*. It fused with more suffixes than in Bonan resulting in the forms listed in Table 5.75. In addition there are also two markers *ba*, *le*, and *sa* that are most likely of Mandarin Chinese origin (e.g., Mandarin *ba* 吧, Xining *le*<sup>53</sup> 呢, Hezhou *za*<sup>3</sup>, see §5.9.2.1) As in Bonan, content questions are usually unmarked.

Table 5.75: Special question forms in Kangjia (Siqinchaoketu 1999: passim; 2002: passim)

Meaning	Declarative	Interrogative
NFUT	-m,	-m <u>u</u>
NFUT	-na, ...	-n <u>u</u>
PST	-va, -pa, -ba	-v <u>u</u>
PROGR	-si, -sina	-s <u>u</u> , -s <u>u</u> n <u>u</u>
PROGR	-dzi(na)	-dzin <u>u</u>
PROGR (-dzi igu >)	-dzigu	-dzigu <u>u</u>
PFV	-sun	-sun <u>u</u> /bu/v <u>u</u>
IPFV	-gu(n)	-gu(n) <u>u</u> /bu
HAB (-si-gu >)	-s(u)gu	-s(u)gu bu
FUT (-gui > -gi)	-gu(n)-i/a	-gu(n) bu

(230) Kangjia

a. *tʃi meda-na ba?*

2SG know-NFUT Q

‘You know, right?’

b. *ajɔ! en.de ma-ge-dzi re-va sa?*

EXCL here how-V-CVB.IPFV come-PST Q

‘Oh! Why have you come here?’

c. *tʃi-ni kʉn le?*

2SG-GEN person Q

‘What about your person?’ (Siqinchaoketu 1999: 215, 222, 217)

Alternative questions have only one marker attached to the first alternative.

(231) Kangjia

a. *tʃi mede-nʉ?*

2SG know-NFUT.Q

‘Do you know?’

b. *te tʃi-ni-gʉ bu?*

that 2SG-GEN-?N Q

‘Is that yours?’

c. *re-vʉ, se re-va?*

come-PST.Q NEG come-PST

‘Has (she) come or not?’ (Siqinchaoketu 2002: 71, 169, 217)

Tag questions in Kangjia take the sentence-final marker *ere ~ are*. Note a formally similar tag question marker *ale* in the Turkic language Tuvan (§5.11.2).

(232) Kangjia

*te kʉn lausu mari, ere?*

that person teacher NEG Q

‘That person isn’t a teacher, right?’ (Siqinchaoketu 1999: 197)

Kangjia has a copula system similar to Bonan (Table 5.76).

Table 5.76: Special interrogative copula forms in Kangjia in analogy to Bonan (Kangjia/Bonan) (Siqinchaoketu 1999: 196f., 216, passim; 2002: passim).

	Subjective	Objective	Interrogative
EXIST	i/wi	va/wa	vʉ/wʉ
COP.EMPH	mbi/mbi	mba/mba	mbʉ/mbʉ

5 Survey of the grammars of questions in Northeast Asia

Similar to Bonan and Kangjia, *Santa* preserves the Mongolic interrogative marker as *-u*, which also fused with the preceding verbal ending or copula (Table 5.77). There likewise does not appear to be a “corrogrative” particle.

Table 5.77: Finite tense aspect markers in *Santa* (Kim 2003: 358; Napoli 2014: 39); in Chaolu Wu (1994c) the interrogative forms are given as *-nu* and *-wo-u*

	Meaning	Declarative	Interrogative
DUR	-IPFV	-ne	-nu
TERM	-PFV	-wo	-wu
PROGR	-CVB.IPFV COP	-zhi wo > -zho	-zhi wu > -zhu

(233) *Santa*

a. *ire-nu?*

come-DUR.Q

‘Will you come?’

b. *kijie-zhi       wu?*

sleep-CVB.IPFV COP.Q

‘Is (s)he sleeping?’

c. *chi khala echi-ne?*

2SG whither go-DUR

‘Where are you going?’ (Kim 2003: 358)

d. *ghoni=chini ire-wu?*

sheep=2SG.GEN come-TERM.Q

‘Have your sheep come back?’ (Todaeva 1959: 284, modified transcription based on Kim 2003)

The form *-mu* found in the following alternative question (234) was not mentioned by Kim but is probably comparable to an identical form in Bonan, the so-called narrative interrogative. In *Santa* as well, only the first alternative receives a question marker.

(234) *Santa*

*bi ənə man-sa       jawu-mu ha man-sa       jawu-nə?*

1SG this direction-ABL go-Q       that direction-ABL go-DUR

‘Will I go this way or that way?’ (Chaolu Wu 1994c: 14)

Todaeva (1959: 295), who did fieldwork among the *Santa* in the middle of the 50s, mentions two additional interrogative particles *la* and *ba*. The latter is clearly of Chinese origin (*ba* 吧, Liu Zhaoxiong 1981: 83).

(235) Santa

a. *hhe shi noghi we-la?*  
 that COP dog COP-?Q

‘Is that a dog?’

b. *e.ne kun e.de ire-ne ba?*  
 this man here come-DUR Q

‘This man is coming, right?’ (Todaeva 1959: 295, modified transcription based on Kim 2003)

Perhaps *la* is a loan from Hezhou Mandarin *la*<sup>3</sup> 啦 (§5.9.2.1). Field (1997: 360) claims that Santa has tag questions that have the form of a regular polar question followed by the irrealis negator *ulīə*, which is a very unexpected construction for a tag question. In fact, an analysis as a negative alternative question in which only the first alternative is overtly marked is more likely. Such a situation can also be found in Karlong Mongghul (see Faehndrich 2007: 221) and Minhe Mangghuer (see below).

(236) Santa

*imani mādziə=nu ulīə?*  
 faith know=Q NEG

‘Do you know the faith or not?’ (Field 1997: 360)

The same construction is also mentioned by Liu Zhaoxiong (1981: 79). A slightly different analysis of the use of negators for question marking has recently been given by Napoli (2014: 41). According to him, there are two negators that can fulfill this function.

Events marked with the non-perfective marker *-ne* can only receive the negator (*u*)*lie*. Since this negator can only be used with events marked with this marker, the finite marker can be dropped. This is not the case with *wuye*, since it can negate events marked with *-wo* and *-zho*. Therefore, in order to specify the tense-aspect of the sentence, the marker is obligatory.

(237) Santa

*chi baza-de echi-wo wuye?*  
 2SG PN-LOC go-TERM NEG

‘Did you go to Linxia or not?’ (Napoli 2014: 41)

However, other sources do not mention the negator *wuye* at all. Santa has a partially productive negative verb *ui-* (Liu Zhaoxiong 1981: 73; cf. S. Kim 2003: 362) that could be the basis for *wuye*. For instance, consider the following example.

(238) Santa

*tʂu-ni ada uai-nu, u-wo?*  
 2SG-GEN father COP-Q NEG-TERM

‘Is your father alive or not?’ (Liu Zhaoxiong 1981: 105, simplified)

**Mongghul** has no “corrogrative” particle in content questions (Stefan Georg p.c. 2015), but preserves the polar question marker. Faehndrich (2007) has collected several descriptions of question marking in Mongghul that exhibit certain differences but usually agree in the presence of three question markers such as neutral *uu*, *nuu* after objective *-a*, and *juu* after subjective *-ii* in Karlong. The descriptions disagree about the analysis of the question markers as particles or suffixes. Here, the original variant has been analyzed as enclitic  $=(y)uu$  and all other forms as suffixes. In Karlong Mongghul the forms are

*nu*; after words ending in the objective suffix *-a*, *ju*; after words ending in the subjective suffix *-i*; and *u*; which is used after words ending in other vowels, including /a/ which is not the objective suffix. Short high vowels are deleted before the interrogative particle *u*: (Faehndrich 2007: 221)

In example (239b) it appears in a focus question and does not stand sentence-finally. It does not, however, attach to the apparent focus in the sentence, which is *cge piisee* ‘big belt’. The situation is thus unlike Middle Mongol. Whether the focus position is sentence initial or preverbal remains unclear, but might be responsible for the sentence-final position of the personal pronoun. But see also §5.11.2 on Turkic languages for second person markers following the question marker.

(239) Mongghul (Karlong)

a. *tci dziehun-la=uu?*

2SG marry-V=Q

‘Are you married?’

b. *cge piisee sge-dz-a-nuu tci?*

big belt see-PERF-OBJ=Q 2SG

‘Have you seen a big belt?’ (Faehndrich 2007: 221, 220)

(240) Mongghul

*qi anji xji-gu-i?*

2SG where go-P.FUT-SUBJ

‘Where are you going?’ (Georg 2003a: 303, shortened)

Åkerman (2012) gives a much more elaborate description of the interrogative verb forms in Mongghul. Similar to other Shirongolic languages, the question marker fused with several verb suffixes and copulas (Table 5.78). There is complex interaction of question marking with the domains of tense, aspect, clause type, and perspective. Similar to Faehndrich’s (2007) claim, the question marker *-nu* only appears after the objective forms. However, the interaction of question markers with the other suffixes is much more complicated. In some cases the question marker simply fused with the suffix, e.g. *-wuu* < *-wa* + *-u*. In other cases, those in which Faehndrich (2007) postulated the question marker *-juu*, the analysis is somewhat unclear.

The question suffix *-niu*, for example, might be analyzable as *-ni-u*, but following Faehndrich an analysis as *-n-iu* is more likely. Only in some cases is there a question marker with a long vowel, e.g. *-m-uu*.

Table 5.78: Special interrogative suffixes and copulas in Mongghul (Åkerman 2012: 13ff.)

Meaning	Declarative	Interrogative
PFV SUBJ	-wa	-wuu
IPFV SUBJ	-nii	-niu
STATE SUBJ	-jii	-jiu
FUT SUBJ	-gu.nii, -gui	-gu.niu, -guu
IPFV NEUT	-ni/-nu	-nu
IPFV NEUT	-m/-n	-muu
FUT NEUT	-m/-n	-muu
COP NEUT	wei	wei-u
PFV OBJ	-jia	-jia-nu
IPFV OBJ	-na	-na-nu
FUT OBJ	-gu.na	-gu.na-nu
COP OBJ	wa ~ ma ~ na ~ la ~ ra	wa-nu ~ ma-nu ~ na-nu ~ la-u ~ ra-nu

## (241) Mongghul

- a. *qi Mongghul uguo mudie-nu?*  
 2SG PN language know-IPFV.Q  
 ‘Do you know (how to speak) Mongghul?’
- b. *qi wazar Minyuen-de sou-niu?*  
 2SG city PN-LOC sit-SUBJ.IPFV.Q  
 ‘Do you live at Minyuan in the city?’
- c. *qi laosi wei-u?*  
 2SG teacher SUBJ.COP-Q  
 ‘Are you a teacher?’
- d. *nie-nu moxi shda-gu.na-nu?*  
 this-ACC read can-OBJ.FUT-Q  
 ‘Can you read this?’
- e. *ngan Zhonggui kun na-nu?*  
 3SG PN person OBJ.COP-Q  
 ‘Is he Chinese?’ (Åkerman 2012: 14)

Alternative questions take two question markers that have to be identical in form.

## (242) Mongghul

- qi niudur xi-gu.niu, malang xi-gu.niu?*  
 2SG today go-FUT.SUBJ.Q tomorrow go-FUT.SUBJ.Q  
 ‘Do you go today or tomorrow?’ (Åkerman 2012: 14)

5 Survey of the grammars of questions in Northeast Asia

Probably the most aberrant Mongolic language with respect to the marking of questions is *Mangghuer*. Instead of a simple particle there is a rather elaborate paradigm of forms which, as in Mongghul, includes the dimension of perspective (Table 5.79), typical for adjacent Tibetic languages (§5.9.2). Nevertheless, the suffixes marking polar questions clearly contain the original interrogative particle.

Table 5.79: Paradigm of question marking in Mangghuer (Slater 2003b: 316; Dixon 2012: 386f.)

		PFV	IPFV	FUT
DECL or CQ	SUBJ	-ba	-la bi	-ni
	OBJ	-jiang	-lang	-kunang
PQ	SUBJ	-bu	-la bi-u	-nu
	OBJ	-jinu	-leinu	-kuninu

(243) Minhe Mangghuer

a. *bi ri-jinu?*

1SG come-PFV.OBJ.Q

‘Did I come?’

b. *qi ri-bu?*

2SG come-PFV.SUBJ.Q

‘Did you come?’

c. *gan ri-jinu?*

3SG come-PFV.OBJ.Q

‘Did (s)he come?’ (Slater 2003a: 198)

There is one example of a negative alternative question in which only the first alternative receives question marking while the second takes a negative marker. In the original, *nu* was written detached from the preceding word.

(244) Minhe Mangghuer

*ta ghula qijige kerli=**nu** lai-kerli?*

2PL two flower want=Q NEG-want

‘Do you want two flowers or not?’ (Zhu Yongzhong et al. 1997: 437)

As expected, there are also special interrogative forms of copulas, as shown in Table 5.80. Again, the original question marker can clearly be recognized, but is not completely analyzable.

Table 5.80: Special interrogative copulas in Mangghuer (Slater 2003b: 318); Slater (2003a: 199) in addition has the variant *meinu*, which is identical to *beinu* in meaning; negative copulas in addition have special attributive forms, SUBJ (*u)gui* and OBJ (*u)guang*, while there are no such special forms for declarative and interrogative copulas

	Declarative	Question	Negative
SUBJ	bi	biu	puzhi
OBJ	bang	beinu	puzhang

(245) Minhe Mangghuer

- a. *bi laoshi meinu?*  
1SG teacher COP.OBJ.Q  
'Am I a teacher?'
- b. *qi laoshi biu?*  
2SG teacher COP.SUBJ.Q  
'Are you a teacher?'
- c. *gan laoshi meinu?*  
3SG teacher COP.OBJ.Q  
'Is (s)he a teacher?' (Slater 2003a: 199)

Content questions do not take the interrogative forms of copulas.

(246) Minhe Mangghuer

- tasi ang=ji-ku-ni bi?*  
2PL where=DIR-IPFV-N COP.SUBJ  
'Where are (all of) you from?' (Chen Zhaojun et al. 2005: 16)

For the Halchighul dialect of Mangghuer Zhaonasitu (1981b: 61) mentions the markers *ba* (Chinese *ba* 吧) and *sa* (perhaps Hezhou *za*<sup>3</sup>) with similar meanings.

At a first glance the situation in Mongghul is very different from the other languages mentioned thus far, but this is partly due to the difference in description. In fact, Mongghul has a strikingly similar system that is given again in Table 5.81, following the analysis by Slater (2003b: 316) and Dixon (2012: 386f.) for Mangghuer. In fact, this new analysis allows us to analyze some of the forms further than would be possible otherwise. The so-called "why-question" marker *-ji* in Mangghuer not shown in Table 5.81 might correspond to *-jii* 'STATE.SUBJ' in Mongghul.

The perspective neutral forms were left aside to make the system more comparable with Mangghuer. The two paradigms show both striking similarities and differences. Altogether, the interaction between the domains of perspective, aspect, and tense is almost identical. In general, however, the Mongghul forms are more readily analyzable. There

5 Survey of the grammars of questions in Northeast Asia

Table 5.81: Paradigm of question marking in Mongghul (Åkerman 2012: 13ff.) in comparison with Mangghuer (Mongghul/Mangghuer)

		PFV	IPFV	FUT
DECL or CQ	SUBJ	-wa/-ba	-nii/- <b>la bi</b>	- <b>gu</b> -nii/-ni
	OBJ	-jia/-jiang	-na/-lang	-gu-na/-kunang
PQ	SUBJ	-wuu/-bu	-niu/- <b>la bi-u</b>	- <b>gu</b> -niu/-nu
	OBJ	-jia-nu/-jinu	-na-nu/-leinu	-gu-na-nu/-kuninu

are slight phonological changes as can be seen from correspondences such as Mangghuer *-ba* and Mongghul *-wa*. Mangghuer has apparently innovative imperfective forms that are a combination of the copula *bi*, interrogative *bi-u* (Mongghul neutral copula *wei*, *wei-u*), and a so-called “imperfective auxiliary linker” *-la* (Slater 2003a: 143) that might correspond to the verbal purposive suffix *-la* in Mongghul often found before auxiliaries (Åkerman 2012: passim). The unexpected objective imperfective forms *-lang* and *-leinu*, corresponding to Mongghul *-na* and *-na-nu*, have been contaminated by *-la* but are preserved in the future. In Mongghul the future forms are still identical to the imperfective forms, except for the future participle marker *-gu* (Georg 2003a: 300). In Mangghuer *-ku* is restricted to the objective forms. This parallel allows an at least historically valid analysis of the two Mangghuer forms into *-ku-nang* (Mongghul *-gu-na*) and *-ku-ni-nu* (Mongghul *-gu-na-nu*).

In sum, for most Mongolic languages the information given for question marking in grammatical descriptions is not sufficient for a full typology. Table 5.82 summarizes the different interrogative marking strategies in Mongolic languages for polar and content questions, exclusively. From Table 5.82 it becomes apparent that the internal diversity of interrogative particles within Mongolic is less pronounced than, for example, Tungusic. In fact, Mongolic languages may be classified into four groups according to their polar question markers. Most languages preserve the original polar question marker =*UU*. Dagur has the form =*yee* instead, Moghol apparently lacks any morphosyntactic question marker, and Buryat, together with Khamnigan Mongol, probably borrowed the marker from an Ewenic (Tungusic) language. Shirongolic also forms a subgroup for itself because in all the languages the question marker fused with other elements, which results in a much more complicated situation. The interrogative marker *-mu* has, according to Sandman (2012: 384), been borrowed from Bonan into the Sinitic language Wutun, but a Turkic origin is more likely (see §§ 5.9.2.1, 5.11.2, 6.1).

Shirongolic languages also have special interrogative forms for copulas that are given in Table 5.83.

Some of the forms are still analyzable into a copula and a question marker, e.g. Mongghul *wei-u*. In other languages such a situation may have existed before phonetic erosion and contraction set in, e.g. Bonan *wi + -u = wu*. Mongghul and Mangghuer have different copula forms depending on the category of perspective, while in some variants of

Table 5.82: Overview of polar and content question markers in Mongolic languages; intonation patterns are excluded

Language	PQ	CQ
Dagur	=yee#	=yee#
Tacheng Dagur	-jA#	-jA#
Moghol	-	-
Khamnigan Mongol	=gu#	bei#
Buryat	=gü#	be ~ =b#
Shineken Buryat	=go#/=gu#/=g	be ~ =b#
Mongolian	=(y)UU# (Table 5.72), =(y)ii#	bwii# ~ bwai#, V-b ~ V-e.b, ? =(y)UU#
Khorchin	=(j)UU#, =(j)ii#, =mu#, ?=me#	=(j)ii#, be#, =me#
Oirat	=(y)UU#, =ii#	=b# ~ =w#
Kalmyk	=u#, =iy# ~ =i#	=b# ~ =w#
Ordos	=UU#	-
Shira Yughur	=(j)uu#	bə#, ?=(j)uu#
Mongghul	=uu, (Tables 5.78, 5.79)	-
Mangghuer	=u, (Tables 5.80, 5.81)	-
Bonan	V-u, (Table 5.73), V-si, ma#	-
Kangjia	V-ᠰ, (Table 5.75), sa#	-, le#
Santa	V-u, (Table 5.77), la#	?

Table 5.83: Special interrogative copulas in Shirongolic languages

	Declarative	Question
Bonan (Wu)	wi SUBJ, wa OBJ, mbi SUBJ, mba OBJ	wu, mbu
Kangjia	i SUBJ, va OBJ, mbi SUBJ, mba OBJ	vᠰ, mbᠰ
Bonan (Fried)	wi SUBJ, wa OBJ, bi SUBJ, ba OBJ	wu(u) SUBJ, wa-u OBJ, bu SUBJ, ba-u OBJ
Santa	wo	wu
Mongghul	wei NEUT, wa, ... OBJ	wei-u NEUT, wa-nu, ... OBJ
Mangghuer	bi SUBJ, bang OBJ	bi-u SUBJ, beinu OBJ

Bonan, Kangjia, and Santa this difference is leveled in the interrogative. Special interrogative forms of copulas are also known from Japonic (Shuri, §5.6.2) and Ainuic (§5.1.2).

### 5.8.3 Interrogatives in Mongolic

There are few good descriptions of interrogatives in Mongolic. Most treatments such as those in Janhunen (2003e) mention only a handful of forms and leave them mostly unanalyzed. Most grammatical descriptions for Mongolic languages also do not mention the syntactic behavior of interrogatives. But they seem to generally remain *in situ* (e.g., Fried 2010: 134; Napoli 2014: 40). Nevertheless, there are quite reliable reconstructions for Proto-Mongolic by Janhunen (2003d: 20) that can serve as a basis for further analysis (Table 5.84, see also Poppe 1955: 229f.).

Table 5.84: Proto-Mongolic reconstructions by Janhunen (2003d: 20) and their modern Mongolian correspondences (Janhunen 2012b: 130ff.; Benjamin Brosig p.c. 2018)

Meaning	Analysis	Proto-Mongolic	Mongolian
who		*ke/n	xen
who PL	PL -d	*ke-d	xed
when		*ke.li	WM keli
how		*ke.r	xer
how many	GEN of *ke-d [?]	*ke.d.ü.n	xedn
how much	ACC of *ke-d [?]	*ke.d.ü.(y)i	xedii
when	LOC *-A	*ke.d.i.x-e > *ke.j.i.x-e [?]	xedzee
where	LOC *-A	*ka.mix-a > *kaxa/n-a [?]	xaan'
what		*yan	WM yan
what		*ya.xu/n	yuun
what kind of		*ya.m.bar > *yamar	yamer
what thing		*ya.xu.ma > *yexüme	youm
to do what	*ki- 'to do'	*ya.xa+ki- > *yaxa- [?]	yaa-
which		*ali/n	alyn

Some developments assumed by Janhunen are marked with a question mark as they are not very plausible. There is evidence of a form \*kamixa 'where' in some languages such as Written Oirat *xamigha(a)*. In Middle Mongol, a form *xamiya* is attested twice (Benjamin Brosig p.c. 2018). Locative interrogatives are not derived from the selective interrogative as in Tungusic but nevertheless display parallels with the demonstratives (Table 5.85). The Turkic language Dolgan has a form *kanna* 'where, whither', and a related form *xanna* is found in Yakut. There are surprisingly similar forms in Mongolic, e.g. Khamnigan Mongol *kaana* or Buryat *xaana* 'where'. But the Yakut and Dolgan forms are contractions of an interrogative that is still analyzable in other Turkic languages including Sarig Yughur *qan-ta* 'which-LOC' (§5.11.3).

Table 5.85: Spatial deictics in Mongolian according to Janhunen (2012b: 131), slightly reduced

	PROX (hearer)	DIST	INT
LOC	naa-n'	tzaa-n'	xaa-(n')
LOC ABL	naa-n-aas	tzaa-n-aas	xaa-n-aas
LAT	naa-sh	tzaa-sh	xaa-sh
PROL	naa-g.oor	tzaa-g.oor	xaa-g.oor

Table 5.86 shows five of the interrogatives that can be found in most modern Mongolic languages.

Table 5.86: Five Proto-Mongolic interrogatives and their modern representatives

	*ken 'who'	*yaxun 'what'	*alin 'which'	*kejixe 'when'	*kaxana 'where'
Dagur	xeng	yoon	aly	xejee	xaan
Mongolian	xen	yuun	alyn	xedzee	xaan'
Buryat	xen	yüün	ali	xezee	xaana
Khamnigan Mongol	ken	yeen	ali	kejje	kaana
Ordos	ken	yüün	ali	kejje	kaa
Written Oirat	ken	you/n	ali	kezee	xamigha(a)
Oirat	ken	yuu/n	äl ~ äl-k	keze	xamaa
Kalmyk	ken	yuun	aly(-k)	kezä	xama, aly-d
Shira Yughur	ken	yima	aali	kejje	xana
Santa	kien	yang	ali	giezhe	khala
Bonan	kang	yang	ane	kece(-)	hala
Kangjia	kɔ	jɔ ~ jaŋ	ani(ye)	gəɖʒe	χana
Huzhu Mongghul	ken	ya(a)n	ali	kijee	an-ji(i)
Minhe Mangghuer	kan	ya, yang	a(yi)ge	kejje	ang(-ji)
Moghol	ken ~ kiyān	emah ~ imas etc.	?	keja	?

According to Janhunen (2003d: 20) the stem \**ke-* originally had the meaning 'who' as well as 'what', which is an unlikely scenario from a cross-linguistic point of view. As has been shown by Cysouw (2005), the only place worldwide where this pattern is not extremely rare or altogether absent is South America.

Proto-Mongolic had two resonances (submorphemes), one in \**k~* that is still present in most Mongolic languages but changed to *x~* in Dagur, Buryat and Mongolian, and one in \**y~* that has survived up to today. Similar changes from \**k~* to > \**x~* can be seen in Turkic

5 Survey of the grammars of questions in Northeast Asia

(e.g., Khakas, §5.11.3) or Tungusic languages (e.g., Nanai, §5.10.3). Only the interrogative \**ali/n* ‘which’ does not fit into either type. All Mongolic languages thus have what has been called K-interrogatives in this study. Furthermore, Mongolic also possesses the KIN-interrogative. Amuric (§5.2.3) and especially Tungusic languages (§5.10.3) exhibit several interrogatives that may have been borrowed from Mongolic.

In the following, I will address interrogatives in individual Mongolic languages in turn. Table 5.87 summarizes the interrogatives found in four descriptions of **Dagur**. The etymology of most of these forms has already been given above.

Table 5.87: Interrogatives in Dagur (Martin 1961: 30f., passim; Zhong Suchun 1982: 52; Chaolu Wu 1996: 22; Tsumagari 2003: 141f.; Yu Wonsoo et al. 2008: 63, passim)

	Martin	Zhong	Chaolu	Tsumagari	Yu et al.
which	ali, alin (attr.)	alj	al <sup>l</sup>	aly ~ alin-	ali ~ aalj
who	hen	xən, anii	xən, aniin	xeng, aniing	anija ~ anja
what	joo	joo	jɔɔ	yoo ~ yoon-	jo, jooke, joon
how many/much		jooke	jɔɔkəə	yookie ~ yiekie	jooked
why	iuuu	juguu		yoondaa, yuguu	jugoo
what (kind of)	iamare ~ iamere	jamər	jamər	yamer	jamər
how many	hede, heden (attr.)	xəpəŋ	xəd	xed	xəd, xədi, xədən (attr.)
when	hejee	xəpʒee	xəpʒəə	xeejee	xəjəə, xəjəər
how	here	xər		xer, yoondaa	xərəə
where	haane	xaanə	xaanə	xaan	xaan
whither			xaidaa	xaidaa	xaan, xandiin
whence					xaanaar

Chaolu Wu (1996) also mentions an interrogative verb *jee-* ‘to do what’. The form *iuuu* ~ *juguu* etc. ‘how, why’ may be cognate with Buryat, Chakhar, Khalkha, and Khamnigan *yaa/g-aad* ‘how, why’, which is a perfective converb form of an interrogative verb that has the form *-g/-AA(r) > \*-g/-AA* in Dagur (Tsumagari 2003: 145f.). The medial *-g-* may either be part of the converb or, less likely, of the verbalizer that has the form *-ge*

in Buryat or Khamnigan. The resonance *\*k~* changed to *\*x~* in Dagur, but not in all dialects. The change did not take place in the Qiqihar dialect, which has forms such as *kuu* ‘person’ or *kər* ‘how’ as opposed to *xuu* and *xər* in other dialects (Ding Danqing 1995: 191). Dagur preserved the original interrogative *xən* ‘who’ but also has an innovative form *anii* that ultimately might be somehow related to *\*ali(n)* ‘which’. Similarly, the two Tungusic contact languages of Dagur, Oroqen and Solon have a form *a(a)wu*, which originally meant ‘which one’ but has extended its meaning and has partly replaced the form *ni(i)* ‘who’ that goes back to Proto-Tungusic (§5.10.3). If true, this could be an instance of a shared grammaticalization. But the exact etymology of the Dagur interrogative is not entirely clear. The suffix in *joon-daa* is likely a dative ending followed by the reflexive marker, i.e. *-d-AA* (Tsumagari 2003: 143). This is also the analysis of the form *joon-d-ee* found among the Dagur interrogative paradigms collected by Martin (1961: 30). Martin also lists a plain dative form *joon-de*, which is likely the source of Nanmu Oroqen *joonde* and could also somehow be connected with Solon *yoodon*.

The interrogatives in *Khamnigan Mongol* and *Buryat* (Table 5.88) are very similar. Khamnigan has a more conservative phonology and preserves the initial *\*k~*, which changed to *x~* in Buryat. Interrogatives with the meaning ‘why’ and ‘how’ are derived with the help of the same case and converbial markers in both languages. Yamakoshi (2007a) mentions a Khamnigan form *kədui cag-* ‘what time’, which is probably a loan translation of Mandarin *jǐ diǎn* (§5.9.3.1). Castrén (1857a) collected several paradigms of interrogatives and demonstratives that are given in a simplified and analyzed version in Table 5.89. The paradigms are clearly identical in their case forms, but the demonstratives take an additional stem augmentation /*n*.

As in Dagur and other languages below the dative case form of the interrogative meaning ‘what’ has acquired the meaning ‘why’. For comparison, Table 5.90 lists Dagur demonstrative and interrogative paradigms, but excludes reflexive case markers. There are some differences in phonology and morphology such as the lack of the /*n* in several Dagur forms. However, given the overall similarity of paradigms, these will not be listed for all languages below.

In *Mongolian* the same change from *\*k~* to *x~* as in Buryat occurred (Table 5.92). According to Mostaert’s account of Ordos (Table 5.93), the initial *k~* is preserved except for *χaa* ‘where’.

There are additional forms such as Khorchin *juv gə-j* ‘what say-CVB.IPFV > why’, which is completely parallel to Manchu *ai se-me*, and a Mongolian interrogative verb *xaa-c-* (< *xaa-oc-*) ‘to go where’. The semantic scope of *yamer* ‘what kind of, how’ suggests a connection with Turkic languages (§5.11.3). Chakhar *gecneeng* has a cognate in Ordos *ce‘ts‘ineen* and Cyrillic Khalkha *xecneen* (Benjamin Brosig p.c. 2016). Instead of *xedii* Khalkha usually has the complex form *xer olon* ‘how much’ (Benjamin Brosig p.c. 2016), which might be a calque of a common European formation transmitted via Russian *kak mnogo/как много*. Georg (2003b: 202) only mentions a few forms for Ordos (*ken* ‘who’, *gecineen* ‘how much’, *kejee* ‘when’, *kaa* ‘where’, *yüü/n* ‘what’, and *yamar* ‘what kind of’). But in his list *kaa* still preserves the initial *k-*.

5 Survey of the grammars of questions in Northeast Asia

Table 5.88: Interrogatives in Buryat (Yamakoshi 2011a: 170; Skribnik 2003: 111) and Khamnigan Mongol (Janhunen 2003b: 92; Yamakoshi 2007a: passim); Buryat also has *xedii-dexi* ‘how manieth’ and *xedii-lüülen* ‘in a group of how many’; some variants were excluded

	Shineken Buryat	Standard Buryat	Khamnigan (Janhunen)	Khamnigan (Yamakoshi)
who	xe/n	xe/n	ken	kən
how many	xedii	xedii.n-		kədui
when	xezee	xezee	kejie	kəzie
how		xer	ker	
where	xaa-(na)	xaa-(na)	kaa-na	kaa-na
whither	xaa-s <sup>j</sup>	xai-sha	kaa-si	kaa-s
whence		xaana-haa		
through where	xaa-g-oor	xaa/g-uur		
what kind of	jamar	yamar	yamar	jamar
what	jun, juu	yüü/n	yee/n ~ yuu/n	jun, joon
why (what-DAT)	juun-de	yüün-de	yeen-du	joon-do
to do what (what-V-)	jaa-	yaa-(ge-)	yaa-(g-)	joo ki-
how, why (IPFV.CVB)	jaa-z <sup>a</sup>	yaa-zha		
how, why (PFV.CVB)		yaa-g-aad	yaa-g-aad	
which	al <sup>j</sup>	ali		ali

Table 5.89: Simplified paradigms of interrogatives and demonstratives in Buryat according to Castrén (1857a: 31ff.); only singular forms and not all variants are shown

	this	that	who	what	which
plain	e.nê	te.rê	ke/n	ju/n	ali/n
ACC	enê/n-i	terê/n-i	ken-i	jun-i	alin-i
DAT	enê/n-de	terê/n-de	ken-de	jun-de	alin-da
ABL	enê/n-ehe	terê/n-ehe	ken-ehe	jun-ehe	alin-aha
INST	enê/n-er	terê/n-er	ken-er	jun-er	alin-ar
COM	enê/n-tei	terê/n-tei	ken-tei	jun-tei	alin-tai

Table 5.90: Paradigms of interrogatives and demonstratives in Dagur according to Martin (1961: 28ff.) in analogy to Table 5.89.

	<b>this</b>	<b>that</b>	<b>who</b>	<b>what</b>	<b>which</b>
plain	e.ne	te.re	he/n	joo	ali
ACC	e.n-ii	te.r-ii	hen-ii	joon-ii	ali-i
DAT	e.n-de	te(.re)/n-de	hen-de	joon-de	alin-de
ABL	e.n-eese	te.r-eese	hen-eese	joon-oose	ali-eese
INST	e.n-eere	te.r-eere	hen-eere	joon-oore	ali-eere
COM	e.n-tei	te.re-tei	hen-tei	joon-tei	ali-tei

Table 5.91: Interrogatives in Mongolian (Janhunen 2012b: 130ff., 255f.) and in Chakhar (Sechenbaatar 2003), Darkhat (Gáspár 2006: 46), and Khorchin dialects (Yamakoshi 2015: passim); not all forms and variants are listed

	<b>Mongolian</b>	<b>Chakhar</b>	<b>Darkhat</b>	<b>Khorchin</b>
who	xe/n	xeng ~ xen-	xen	xən
how many	xed//n	xed, xede-ng (attr.)		xədən
how much	xedii	xedii		xədīi, xədəə
when	xedzee	xejee, xediis	xejee	xəjəə
how	xer			
where	xaa(-n')	xaa(-na=n)	xaa(-nǎ)	xaa
what	yuu/n	yuu, yuu/n	yuu	jʊʊ
why (DAT)	yuun-d		yuun-d	
how, by what means (INST)	yuu-geer		yuu-gaar	
what kind of, how	yamer	yamar	yamǎr	jamar
to do what	yaa-, yuu xii-		yaa-	jaa(-x)-, jʊʊ xii-
how (CVB.IPFV)	yaa-j		yaa-ǰ	jaa-ǰ
why (CVB.PFV)	yaa-gaad		yaa-gaad	
how much/ many		gecneeng		
which	aly//n	aly ~ äly/n ~ älya/n	äli	æl'

5 Survey of the grammars of questions in Northeast Asia

Similar to Ordos, in both **Oirat** and **Kalmyk** the \**k*- remained stable in the stem \**ke*- but changed to *x* in the stem \**kaa*- (Table 5.92). The same is true for Shira Yughur and maybe for Santa as well (see below). The form *aly-d* ‘where’ in Kalmyk clearly is a locative (dative) form of the interrogative *aly* ‘which’. Spoken Oirat *äl-k* and Kalmyk *aly-k* in all likelihood have the same origin as Mangghuer *ali-ge*. Instead of Kalmyk *xamaran* ‘whither’ my informant employed the form *al’daran*, based on *aly* ‘which’.

Table 5.92: Interrogatives in Oirat and Kalmyk (Birtalan 2003: 220; Bläsing 2003: 239)

	Written Oirat	Spoken Oirat	Kalmyk
who	ken	ken	ken
when	kezee	keze	kezä
how much/many	kedüi ~ kedüü	kedn ~ kedü	kedü ~ kedü/n-
what	you/n	yuun/n	yuun
to do what			yagh-
why (-ad CVB.PFV)			yagh-ad
what kind of	yamaaru	yamr/n	yamr ~ yamaran
where (-d LOC)	xamigha ~ xamighaa	xamaa	xama, aly-d
whither			xamaran
which	ali	äl ~ äl-k	aly ~ aly-k

In **Shira Yughur**, however, roughly half of the interrogatives show a resonance in *y*-. Most interrogatives are either inherited from Proto-Mongolic or have a straightforward explanation such as a contraction with a following verb or the presence of a case marker. Only the form *yima* ‘what’ clearly differs from Mongolian. Its explanation is probably related to the change of meaning of the interrogative *yaan* from ‘what’ to ‘how’.

The **Bonan** interrogative *χala* ‘where’ similar to Santa *khala* has a liquid *l* instead of a nasal *n* (cf. Mongolian *xaan*). Whether *anə* ‘which’ is cognate with Mongolian *alyn* ‘which’ or rather Dagur *anii* ‘who’ remains unclear to me. The forms *jant<sup>h</sup>oχ* and *yamtig* from the two different descriptions probably represent dialectal variants of one and the same interrogative. The interrogative *yamten’ge* ‘how much’ is said to result from a fusion with the numeral *nege* ‘one’. This development of the numeral ‘one’ appears to have been influenced by Tibetic (§§3.5, 5.9.3.2). Both ‘how’ and ‘why’ are clearly based on the interrogative verb *jan-gə-*, which in turn is transparently derived from *jan-* ‘what’.

**Kangjia** has the same derivation *jan-gi-* ‘to do what’ but the stem *jan* ‘what’, possibly in analogy to *kə* ‘who’, also has the alternative form *jə* ‘what’. As opposed to Bonan, but similar to Santa, there is an interrogative stem *ma-*. As opposed to Bonan *yamten’ge* ‘how much’, Kangjia has *ma-tu niye* that is derived from *matu* but is likewise based on the numeral ‘one’.

The origin of the **Santa** interrogative *dzidziən-də* is Chinese *jǐ-diǎn* ‘what time’. But it contains an autochthonous dative (locative) marker *-də* that is also present in the

Table 5.93: Interrogatives in Shira Yughur (Nugteren 2003: 273; Zhaonasitu 1981a: 27, *passim*), and Ordos (Mostaert 1937: *passim*)

	Shira Yughur		Ordos
	Nugteren	Zhaonasitu	
who	ken	ken	ke'n
what			juu, juun
how much (niye 'one')	keedi	keedə niye	ke'd <sup>u</sup> ii, ge'ts'ineen
how many	kedən	keedə	ke'du
when	kejee	kedzee	ke'džee
where	xana	χana	χaa(-na)
whither	xagshi	χaɕʃə	
whence	xanasa		
to go where (*kaana od- >)	xanad-		
how (INST)	yaan		juu-gaar
why (DAT, CVB.PFV)	yaan-di	jaan-də	juun-du, juundaan, jaax <sup>k</sup> χuun
what	yima	ima	
what kind of, how	yimar	imar	jamar
to do how	yaa-gi-	ima-γə-	
to do what	yima-gi-		jaa-(χ <sup>k</sup> χi-)
to happen how (*yama bol-)	yimal-		
which	aali	aalə	ali

complex expression *ali orŋ-də*, which literally means 'at which place' and has parallels in several languages such as Mandarin Chinese *zài shénme dìfang* '(COP.)LOC what place' but also Kangjia *ani satfa*. Another loan from Mandarin is the second part of *yan shihou* 'what time' (Mandarin *shìhòu* 'time') that is also present in Kangjia *ani-ye sux-du*.

The form *matu* 'how' looks very untypical for Mongolic. According to Siqinchaoketu (1999: 194), Kangjia *ma-* is an abbreviated form of *jama*, which seems possible but is in need of further explanation. A more plausible alternative would be a Sinitic origin, e.g. Mandarin *mà* (§5.9.3.1). The second part of *ma-tu* could be a derivational suffix that attaches to nouns to form adjectives (Kim 2003: 352). It may be noted that in Kangjia the suffix *-tu* is optional in the interrogative verb *ma(-tu)-gi-* 'to do how'. Interestingly, in Santa *matu* is usually followed by the verb *gie-* 'to say, to make, to think' unless it has the form *ma-tu-kaŋ*. In fact, Santa *ma-tu gie-* looks suspiciously similar to Kangjia *ma-tu-gi-*. The suffix *-kaŋ* possibly derives nouns from adjectives (*-ghang* in Kim 2003: 352), which would explain why it is followed by a copula in the following example.

5 Survey of the grammars of questions in Northeast Asia

Table 5.94: Interrogatives in Bonan (Fried 2010: 144, 261; Hugjiltu 2003: 337, 342) and Kangjia (Siqinchaoketu 1999: 185ff., passim; 2002: 73); see also Todaeva (1963: 187)

	Bonan		Kangjia
	Fried	Hugjiltu	
who	k <sup>h</sup> aŋ	kang	kə
whose (-gaŋ POSS)	k <sup>h</sup> aŋ-gaŋ	kang-g(h)ang	
how many	k <sup>h</sup> ət <sup>h</sup> oŋ	kudung ~ kutung	gədo ~ gədo
which (niye ‘one’)	anə	ane	ani, aniye
when (-də DAT)			aniye suχəu-də
where (satʃa ‘place’)			ani satʃa
what thing			jama(-sə(n))
where	χala	hala	χana
whence			χana-sa(la)
when	k <sup>h</sup> ətɕə ~ k <sup>h</sup> ətɕə-χaŋ	kece ~ kece- xangnang	gədʒe
what (=gə SG.IND)	jaŋ ~ jaŋ-gə	yang	jə ~ jaŋ
to do what (-gə/-ge/-gi V)	jaŋ-gə-	yang-ge-	jaŋ-gi-
how (-tɕə/-je CVB.IPFV)	jaŋ-gə-tɕə	yang-ge-je	
why (-saŋ P.PFV, ?-da DAT)	jaŋ-gə-saŋ	yang-ge-da	
what kind of	jaŋ <sup>h</sup> oχ	yamtig	
how much (nege/niye ‘one’)		yamten’ge	ma-tu niye
how (-tu ADJ)			ma-tu
to do how (-gi V)			ma-tu-gi- ~ ma-gi-

(247) Santa

- a. *tci ma.tu gie-wo?*  
 2SG how do-TERM  
 ‘How are you doing?’
- b. *ən.udu.ku tciəntci ma.tu.kəŋ wo?*  
 today weather how COP  
 ‘How is the weather today?’ (Chaolu Wu 1994c: 17)

In Todaeva (1959: 288) we find an additional form *ma.tu.n-ni* ‘what kind of’ with a third person possessive ending. The complex interrogative *yan gie-zhi* ‘why, what for’ can be analyzed as ‘what do/say-CVB.IPFV’ and is a parallel to Bonan *yang-ge-je*. Not mentioned in Table 5.95 are plural forms such as *jan-la* (Ma Guoliang & Liu Zhaoxiong 1986: 174), which carries the special Santa plural marker of unclear origin. The initial \**k* has three or four different reflexes (*k*, *g*, and *q* ~ *kh*). It may be noted that today in both Santa and

Table 5.95: Interrogatives in Santa (Chaolu Wu 1994c: passim; Kim 2003: 356; Napoli 2014: 40); see also Field (1997: 360); not all variants are listed

	Chaolu Wu	Kim	Napoli
who	kiən	kien	kien
how much/many	giədun, giədu-kaŋ	giedun	giedun
when		giezhe	giezhe
where	qala	khala	khala
whence			khala-se
what	jaŋ, ja	yang	yan
why (gie-zhi ‘do/say-CVB.IPFV’)			yan gie-zhi
which	ali	ali	ali
where (-də DAT)	ali oŋ-də		
how (gie- ‘to do/say’)	matu gie-, matu-kaŋ		matu gie-
when, what time	dzidziən-də		yan shihou

Kangjia the interrogative ‘who’ is the only interrogative starting with a *k-*, for which there may well be functional rather than phonological reasons (§6).

The majority of interrogatives in **Mongghul** start with *a~*, several with *k~* and only one or two with *y~* (Table 5.96). Faehndrich (2007: 127) mentions one form *ticaan* ‘how many’ that seems to have been borrowed from an unknown source. For the Halchighul dialect of Mongghul, Schröder (1964: 151) lists the interrogatives *kän* ‘who’, *yan* ‘what’, *ali* ‘which’, and *kidi* ‘how much’.

Table 5.96: Interrogatives in Huzhu Mongghul (Chaolu Wu 1994b: passim; Faehndrich 2007: 127; Dpal-ldan-bkra-shis et al. 1996: passim)

	Chaolu Wu	Faehndrich	Dpal-ldan-bkra-shis
who	ken	kani SUBJ, kana OBJ	ken
how much/many		kidi, ticaan	kidi-hangi
when (-di DAT)	kədzee	kčee, ali-sxuu-di	kijee, ali sghuu
which	alə		ali
which one			alinga, alingi
what kind of			amahgi sanba
how, why	amaga, amarr-	amagidza	amaga, amakiji
where	andzii	andzii SUBJ, andza OBJ	anji(i)
what	jaan	jaanii SUBJ, jaana OBJ	yan, yanna

5 Survey of the grammars of questions in Northeast Asia

The second part of *amahgi sanba* ‘what kind of’ mentioned by Dpal-Idan-bkra-shis et al. (1996: 232, 241) means ‘kind, type, pattern’. The form *ali sghuu/ali-sxuu-di* literally means ‘at which time’. A speciality of Mongghul is the presence of a perspective distinction in several interrogatives, which was only mentioned by Faehndrich (2007).

Table 5.97: Interrogatives in Minhe Mangghuer (Slater 2003a: 55, 86; Dpal-Idan-bkra-shis et al. 1996: passim)

	Slater	Dpal-Idan-bkra-shis
who	kan	kan
how much/many		kedu
when		kejie
why, how (=la INST, =ji DIR)	ya=la, ya=ji, ya	ya-la, ya-ge, ya-ji
what	yang, ya	yang
which (one)	ayi-ge	ali, a-ge, ali-ge
what kind of	amer-da	yamer(-da)
where (from yang?)	ang	ang-ji

Neither Slater (2003a,b), nor Dpal-Idan-bkra-shis et al. (1996) give a clear analysis of the Mangghuer interrogatives (Table 5.97). But *kan*, *kedu*, *kejie*, *yang*, *ali*, and *yamer* are clearly of Proto-Mongolic origin. The form *angji* ‘where’, also present in Mongghul as *anjii*, probably contains a case ending that was given as a directive =*ji* by Slater (2003b: 312) and is specific to Mangghuer. Problematically, it expresses only direction but not location, for which there is the dative/locative =*du*. The form *ya=ji* ‘why’ thus literally means ‘where to’ (cf. English *to what end*). The comparison of the two forms *amerda* and *yamerda* (both with an unclear suffix *-da*) with and without initial approximant suggest that the form *ang* might be a variant of the interrogative *yang* ‘what’. The interrogative *ayige* or *age* probably contains the indefinite singular marker =*ge*, which is either derived from the Chinese classifier *ge*  $\hat{u}$  (via the loan *yige* ‘one’, from *yi-ge* ‘one-CLF’), or from the autochthonous numeral *nige* ‘one’ (cf. Slater 2003a: 100). This analysis is corroborated by the form *ali-ge* ‘which one’. But the first part *ayi-* or *a-* remains unclear from a language-internal perspective. Most likely it has been borrowed from a Sinitic language (§5.9.3.1). The corresponding form in Mandarin is *nǎ(-yi)-ge* with or without the numeral ‘one’. This may explain the difference between *ayi-ge* and *a-ge*. In Sinitic languages of the area the interrogative lost the initial nasal, e.g. Hezhou/Linxia *a-ʒi<sup>24</sup>-gə* ~ *a-ji<sup>24</sup>-gə*.

## 5.9 Trans-Himalayan

### 5.9.1 Classification of Trans-Himalayan

This study includes languages from three of the 42 subbranches of Trans-Himalayan (van Driem 2014: 10). These branches are Sinitic, Bodish, and Qiangic. Of Bodish, only the Tibetic subbranch will be included here. *Glottolog* (Hammarström et al. 2016) mentions 475 Trans-Himalayan languages, which are almost all located to the south of NEA. The exact relation of the individual subbranches is somewhat disputed and is of no particular concern here. Despite certain controversies (Kurpaska 2010: 25–62), Sinitic is usually divided into seven often mutually incomprehensible main dialect areas called Gan (gàn 贛), Hakka (*kèjiā* 客家), Mandarin (*guān* 官), Min (*mǐn* 闽), Wu (*wú* 吴), Xiang (*xiāng* 湘), and Yue (*yuè* 粤). The existence of separate Pinghua (*pínghuà* 评话), Jin (*jīn* 晋), and Hui (*huī* 徽) dialects is somewhat disputed. Of these, Mandarin is the largest and, if one includes Jin and Hui within it, the only one located in NEA. Recent migrations of speakers of other dialects will not be considered. Mandarin itself may be classified into several regional varieties, of which the Southeastern area in and around Sichuan, Chongqing, Guizhou, and Yunnan as well as the Jianghuai subdialect area around the lower Yangtze are mostly excluded from this study. An exhaustive overview of questions in all the remaining Mandarin subdialects is impossible to give because of a lack of high quality materials. The focus will lie on a description of Standard Mandarin, to which will be added a sample of regional dialects for which good data was available, especially for interrogatives. Special cases to be addressed are Dungan (*dōnggān* 东干, *dunganskij*/дунганский), Gangou (*gāngōu* 甘沟), Hezhou (*hézhōu* 河州) or Linxia (*línxià* 临夏), Tangwang (*tāngwāng* 唐汪), and Wutun (*wūtún* 五屯). With the exception of Dungan, spoken in several Central Asian countries, but derived from Northwest China, these languages are all spoken in the Qinghai-Gansu area and exhibit a strong influence from Turkic, Mongolic, or Tibetic.

Tibetic alone encompasses about 200 different varieties, divided into eight different groups or sections (Tournadre 2005; 2014). Here only Amdo Tibetan (*ānduō* 安多) dialects and gSerpa (*sè'ěrba* 色尔坝) from the northeastern, as well as Baima (*báimǎ* 白马), Cone (Chone, *zhuōní* 卓尼), and Zhongu from the eastern section will be included. I currently lack sufficient data for some varieties such as Khalong Tibetan from the eastern section spoken in northern Sichuan (see Sun 2007).<sup>22</sup> Both Zhongu and Baima are sometimes considered Qiangic instead of Tibetic. Amdo Tibetan is one of the more dominant languages of the Amdo Sprachbund (Sandman & Simon 2016) and is said to have at least 23 different subdialects (Ebihara 2011: 43). Given the limited amount of information available, only a fraction of this variation can be included here. Of the disputed Qiangic branch of Trans-Himalayan, only the extinct Tangut language, the major language of the Xixia empire (1038-1227), was located in NEA. The language was only rediscovered and deciphered in the 20<sup>th</sup> century (Gong Hwang-Cherng 2003).

Descriptions of Chinese dialects frequently suffer from a lack of accuracy and analysis, the use of characters for transcription, and an incoherent use of characters for etymolog-

<sup>22</sup>There seems to be no official Chinese name for Zhongu or Khalong yet.

ical and phonological purposes. This unfortunate and confusing situation could only be partly remedied here. Chinese dialectal or historical data that was available in Chinese characters exclusively will be transcribed with the official romanization system Pinyin without tones in square brackets.

## 5.9.2 Question marking in Trans-Himalayan

### 5.9.2.1 Question marking in Sinitic

Old Chinese had a sentence-final polar question marker 乎 [*hu*] (see 248) that has been reconstructed as \**ɣʰa* by Baxter & Sagart (2014b).

- (248) Old Chinese  
賢者亦樂此乎  
[*xian zhe yi le ci hu*]  
virtuous N also enjoy this Q  
‘Does a man of virtue also enjoy such (things)?’ (Pulleyblank 1995: 139)

Its form is somewhat reminiscent of interrogatives that will be described in the next section. There are several other question markers that appear to be contractions of [*hu*] 乎 with other elements and will not be discussed any further here (see Pulleyblank 1995: 139ff.). Content questions were generally unmarked.

- (249) Old Chinese  
何必曰利  
[*he bi yue li*]  
why must say profit  
‘Why must you say ‘profit’?’ (Pulleyblank 1995: 145)

Of course, there is a lot of variation to be found in historical stages of Chinese, but a detailed examination of diachronic developments that necessarily also includes all modern Sinitic languages goes well beyond the possibilities of this study. Instead, the following will address exclusively the modern Sinitic languages located in NEA.

**Standard Mandarin** Chinese data are partly based on my own knowledge and were partly elicited or confirmed in 2015 and 2016 with the help of a native speaker from Guiyang, fluent in both the dialect and standard Mandarin, living in Germany. Mandarin Chinese has many different question marking strategies, but the default form is the sentence-final marker *ma* 吗 that marks polar questions.

- (250) Mandarin  
*chī-fàn le ma?*  
eat-food PFV Q  
‘Have you eaten yet?’

Polar questions may also simply be marked with rising intonation, though this seems less common. Content questions are usually unmarked morphosyntactically and have an intonation contour similar to a declarative sentence, but with additional emphasis of the interrogative.

- (251) Mandarin  
*nǐ jiào shénme?*  
 2SG call what  
 ‘What is your name?’

The marker *ba* 吧 has a function similar to a polar question and for practical purposes is classified as such. But it has an additional element of supposition by the speaker. In some instances it is best translated as a tag question in English.

- (252) Mandarin  
*nǐ chī-fàn le ba?*  
 2SG eat-food PFV Q  
 ‘You must have already eaten (I suppose)?’

In another function, *ba* 吧 is also an imperative marker. In its function as question marker, which is similar to English *must* in relatively certain hypotheses, it has been adopted by a great many languages in China, perhaps because of its very specific semantic nuances. Both *ma* and *ba* can be the actual question marker in complex tag questions.

- (253) Mandarin
- a. *nǐ chī-fàn le, duì ba?*  
 2SG eat-food PFV right Q  
 ‘You have already eaten, right?’
  - b. *nǐ bāng wǒ mǎi.dōngxī, hǎo ba?*  
 2SG help 1SG shop good Q  
 ‘You are going shopping for me, right?’

In both cases *ba* may be substituted with the more neutral *ma*, which is accompanied with a slight change in meaning. Several more patterns are possible, e.g. *shì ma?* 是吗 ‘cop Q’, *kěyǐ ma?* 可以吗 ‘be.possible Q’, *xíng ma?* 行吗 ‘be.possible Q’.

Mandarin has a further colloquial marker *ne* 呢 that has both an interrogative and non-interrogative function (Li & Thompson 1981: 300–307). In its interrogative function it has an interesting distribution. It can be found in negative alternative questions (A-not-A), content questions, and “truncated” or elliptical “questions consisting of only one noun” (Li & Thompson 1981: 305) such as *tā ne?* ‘(And) how about him/her?’.

- (254) Mandarin
- a. *nǐ qù bu qù zhōngguó ne?*  
 2SG GO NEG GO PN Q  
 ‘Are you (actually) going to China?’

- b. *nà nǐ qù nǎli ne?*  
 that 2SG go where Q  
 ‘Well, where are you going then?’

The last function (254b) is what might be called a topic question, the exact meaning of which depends on the previous discourse. Yet another connection of *ne* 呢 with questions is a special type of alternative question construction that we will encounter further below.

As for focus questions, there are several possible patterns. One is a cleft-like structure with the copula *shì* 是 and occurs before the focused element. Compare the following examples of a polar and two focus questions. If the copula stands sentence-finally, it functions as a question tag.

(255) Mandarin

- a. *nǐ qù zhōngguó, shì ma?*  
 2SG go PN COP Q  
 ‘You are going to China, are you?’
- b. *nǐ shì qù zhōngguó ma?*  
 2SG COP go PN Q  
 ‘Is it *China* that you are going to?’
- c. *shì nǐ qù zhōngguó ma?*  
 COP 2SG go PN Q  
 ‘Is it *you* who is going to China?’

The same sentences are possible with the marker *ba* 吧. Another marker for focus questions is likewise based on the copula but has itself a structure of an A-not-A question, *shì-bu-shì* 是不是 ‘COP-NEG-COP’, which is why no additional question marker is present. It is treated as a single marker here that stands before the focused element or attaches to the sentence and functions as a question tag.

(256) Mandarin

- a. *nǐ qù zhōngguó, shìbushi?*  
 2SG go PN Q  
 ‘You are going to China, aren’t you?’
- b. *nǐ shìbushi qù zhōngguó?*  
 2SG Q go PN  
 ‘Is it *China* that you are going to?’
- c. *shìbushi nǐ qù zhōngguó?*  
 Q 2SG go PN  
 ‘Is it *you* who is going to China?’

In the latter two examples (256b, 256c), *shì-bu-shì* may be replaced with the more literary *shì-fǒu* 是否, a combination of the copula with an otherwise uncommon negator. This is part of the literary language and cannot function as a question tag. But there are several related question tags such as *duì-bu-duì* 对不对 ‘correct-NEG-correct’, *hǎo-bu-hǎo* 好不好 ‘good-NEG-good’ *xíng-bu-xíng* 行不行 ‘be.possible-NEG-be.possible’, *kě(yì)-bu-kě(yì)* 可以不可以 ‘be.possible-NEG-be.possible’, or *huì-bu-huì* 会不会 ‘be.able-NEG-be.able’, all of which are of the A-not-A type but do not usually mark focus questions. This is evidence that *shì-bu-shì* actually has the status of a sentence (A-not-A question) in tag questions, but of a question marker in polar and focus questions.

Alternative questions have two main construction types, either mere juxtaposition or the use of an interrogative disjunctive *háishì* 还是 (different from the standard disjunctive *huòzhě* 或者) (cf. Hölzl 2016b: 20).

(257) Mandarin

- a. *zhēn.de* Ø *jiǎ.de*?  
 true (or) false  
 ‘Is that true or false?’
- b. *nǐ qù zhōngguó háishì rìběn*?  
 2SG go PN or.Q Japan  
 ‘Are you going to China or Japan?’

One very dominant question category in Mandarin Chinese are negative alternative questions that exhibit the same two marking strategies as plain alternative questions. Juxtaposition is much more frequent and productive in negative alternative questions (A-not-A questions) than in plain alternative questions (Hölzl 2016b: 20).

(258) Mandarin

- nǐ qù zhōngguó (háishì) bú qù zhōngguó*?  
 2SG go PN or.Q NEG go PN  
 ‘Are you going to China or are you not going to China?’

In some cases the whole second alternative is deleted except for negation. This may be the basis for the grammaticalization of interrogative particles such as *ma* 吗. Note that the following two sentences are completely identical in structure. The major difference appears to be the fact that *ma* is restricted to this construction while *bù* 不 is still a productive negative marker otherwise. But note that in this context sometimes it had already lost its tone, which is a sign of grammaticalization (i.e., phonetic erosion, Hölzl 2015e).

- (259) a. Mandarin  
*nǐ qù zhōngguó bu*?  
 2SG go PN NEG

- b. *nǐ qù zhōngguó ma?*  
 2SG go PN Q  
 ‘Are you going to China?’

In the past tense or in sentences that contain the existential *yǒu* 有, it is also possible to replace the second alternative with the negative existential. In the former, case an additional marker in the first alternative is necessary. Following Sun Chaofen (2006: 64–70), these may glossed as experiential (EXP) *guo* 过 and perfective (PFV) *le* 了, respectively.

- (260) Mandarin  
*nǐ qù-guò/le zhōngguó méi.yǒu?*  
 ’2sg go-EXP/PFV PN NEG  
 ‘Have you been to China or not?’

My informant tells me that, in the first case, the number of times one has been to China as well as the exact time is irrelevant, while in the latter the event is thought to have happened once and relatively recently. Note that the second alternative may not consist of the disjunction and a negator, exclusively (*\*háishi bu?*) (Luo Tianhua 2013: 186), which represents a difference with respect to MSEA (Clark 1985).

In extreme cases the whole second alternative is deleted and alternativity is indicated with the help of the disjunctive connective *háishi* 还是 ‘or.Q’, exclusively. These developments crucially depend on the context of an elliptical (more precisely anaphoric) alternative question. Mandarin also allows open alternative questions.

- (261) Mandarin  
*nǐ qù zhōngguó háishi nǎ.r?*  
 2SG go PN or.Q where

In this case, my informant tells me, one of the interrogatives *nǎ* 哪/*nǎ.r* 哪儿/*nǎ.li* 哪里 ‘where’ would seem more natural than *shénme* 什么 ‘what’, which, however, is possible in other examples.

An element that can often be encountered in Mandarin questions is the sentence-final marker *a* 啊 ~ *ya* 呀 that I analyze as enclitic =(y)a. It is not a question marker as such but “has the semantic effect of softening the query” (Li & Thompson 1981: 313). Some examples will be given further below. Following these authors, the enclitic will be glossed as reduced forcefulness (RF).

As we have just seen, question marking in Mandarin is relatively complex and exhibits many different constructional patterns. The same is probably true for the other Sinitic languages surveyed in the following. But in the absence of native speakers and a lack of detailed information, only some limited information can be given here for each language. Chinese as spoken by the Hui (Chinese speaking Muslims) in **Urumqi** is relatively close to Standard Mandarin. Polar questions have a cognate of *ma* 吗 and content questions optionally have a cognate of *ne* 呢. As expected there are also negative alternative questions.

(262) Mandarin (Urumqi Hui)

- a.  $tʂəŋ^{24}.ti^{21} ma^{21}?$   
 real Q  
 ‘Really?’
- b.  $ni^{52} ʂā^{44} nər^{24} tɕ^{44}?$   
 2SG DIR where go  
 ‘Where are you going?’
- c.  $sei^{24} tciʋu^{44} ʋɿ^{52} nə^{21}?$   
 who save 1SG Q  
 ‘Who will save me (then)?’
- d.  $ni^{52} nəŋ^{24} pu^{21} nəŋ^{24} lɛ^{24}?$   
 2SG be.able NEG be.able come  
 ‘Are you able to come?’ (Liu Liji 1989: 222, 219, 206, 217)

Similar to Standard Mandarin, question tags may contain a question marker, e.g.  $xɔ^{52} pa^{21}$  好吧 or may have the form of an A-not-A question.

(263) Mandarin (Urumqi Hui)

- $tʂ^{21} liɔ^{24} fǎ^{44} tɕ^{44} tɕ^{44}, xɔ^{52} pu^{21} xɔ^{52}?$   
 eat PFV meal again go good NEG good  
 ‘Let’s go after having eaten, alright?’ (Liu Liji 1989: 221)

In the following example of an alternative question (264), the first alternative receives the marker *ne* 呢 that combines with the disjunction and necessarily is preceded by the copula that precedes the first focused element.

(264) Mandarin (Urumqi Hui)

- $ni^{52} sɿ^{21} tʂ^{44} ʋu^{24} jǎ^{21} ni^{44} xɛ^{24} sɿ^{21} xɿ^{21} ts^{44}?$   
 2SG COP smoke Q or drink tea  
 ‘Do you (want to) smoke or drink tea?’ (Liu Liji 1989: 221)

This has an exact parallel in Standard Mandarin (*nǐ shì chōuyān ne háishì hē chá?*).

An idiosyncratic pattern is the presence of a question marker on the first alternative in alternative questions that may also lack a cognate of Mandarin *háishì* 还是. This pattern was probably influenced by surrounding Turkic languages. Especially intriguing is the optional combination of two markers, which is impossible in Standard Mandarin but can also be found in Hezhou Chinese (see Table 5.98 below).

(265) Mandarin (Urumqi Hui)

- $ni^{52} ta^{44} ni^{21} ma^{21} ʋɿ^{52} ta^{44}?$   
 2SG big Q Q 1SG big  
 ‘Are you older or I?’ (Liu Liji 1989: 211)

Xining Mandarin has unmarked content questions, though cognates of Mandarin  $=(y)a$  啊/呀 and  $ne$  呢 are optionally present. Polar questions have a marker  $m\omega^{53}$  which does not appear to be a cognate of Mandarin  $ma$  吗, which is attested as  $ma$  (Zhang Chengzai 1980: 300). The difference with Standard Mandarin is mostly phonological in nature. The formal similarity to the Tangut sentence-final question marker  $mo^2$  is probably accidental.

(266) Xining Mandarin

a.  $f\epsilon i^{2421} ia^{4453}?$

who RF

‘Who (is it)?’

b.  $l\omega^{53} s\gamma^{21321} l\epsilon^{53}?$

HON four Q

‘What about Laosi (the fourth brother)?’

c.  $teia^{44} xa^{35} m\omega^{35} f\omega^{44} u\tilde{a}^{3521} m\omega^{53}?$

3SG yet NEG speak finish Q

‘Has (s)he still not finished speaking?’ (Zhang Chengzai 1980: 300)

There is also the development from negation to question markers. In the following example Standard Mandarin employs the affirmative potential marker (*ná-de-dòng*).

(267) Xining Mandarin

a.  $na^{24} tu\grave{a}^{213} lia^{53} (p\gamma^{44})?$

take move ?POT NEG

‘Can you take it?’

b.  $ni^{53} f\tilde{a}^{213} t\zeta^{44} li\omega^1 m\omega^{24}?$

2SG meal eat PFV NEG

‘Have you eaten?’ (Zhang Chengzai 1980: 301)

The negators are cognates of Mandarin *bù* 不 (non-past) and *méi* 没 (past). A native speaker living in Germany in January 2017 made me aware of the fact that the use of  $m\omega^{24}$  in example (267a) is not only possible, but perhaps more natural. In April 2017 the following examples of alternative and focus questions were recorded. The transcription and analysis roughly follow Zhang Chengzai (1980). Given that the speaker appears to show strong influence from Standard Mandarin, tones were omitted for simplicity.

(268) Xining Mandarin

a.  $ni\ t\zeta\eta\eta ku\ ts'ɿ-lia\ ma\ xais\lbracket\ zɿp\grave{a}n\ ts'ɿ-lia?$

2SG PN go-?POT Q or.Q PN go-?POT

‘Are you going to China or to Japan?’

- b. *tʃuŋku te 'y tʃuɔtsɿ ʂl̩ ni sa?*  
 PN go ?N COP 2SG Q  
 'Is it you who is going to China?'

Example (268a) apparently contains cognates of Mandarin *ma* 吗 and *háishi* 还是. The question marker *sa* in example (268b) appears to also exist in Hezhou Chinese (see below).

**Dungan** questions appear to be very close to Mandarin as well. In the first two examples the original was in traditional characters that have been changed into simplified characters. In the last example the Chinese characters have been added by me. Dungan is usually written with Cyrillic letters, however.

(269) Dungan

- a. 你们欧洲东干人多吗?  
*[ni-men ouzhou donggan ren duo ma?]*  
 2SG-PL PN PN person many Q  
 'Are there many Dungan in Europe?'
- b. 咋的呢?  
*[za-de ne?]*  
 how-ATTR Q  
 'How are you?' (Rimsky-Korsakoff 1994: 486, 515)
- c. [第三段儿话里头狒的啥?]  
*ti san tua.r xua li.t'i fe-ti sa?*  
 ORD three CLF speech inside say-ADV what  
 'What is said in the third sentence?' (Rimsky-Korsakoff 1967: 382)<sup>23</sup>

Question marking in **Hezhou/Linxia** Chinese is very complex and deviates strongly from Standard Mandarin. Table 5.98 summarizes the specialized description of question markers by Xie Xiaolan & Zhang Shumin (1990). Especially interesting is a functional differentiation of three different question markers for polar and content questions each. The marker *mu*<sup>3</sup> most likely was borrowed from Uyghur =*mu* (§5.11). The markers *la*<sup>3</sup> and *za*<sup>3</sup> apparently found their way into some Mongolic languages (§5.8). The combination of two question markers such as *ni*<sup>3</sup>*mu*<sup>3</sup> is similar to Urumqi Hui Chinese. A double marking pattern for alternative questions most likely has been adopted from Turkic as well.

The following two examples of a polar question with the Uyghur question marker (270a) as well as an unmarked content question (270b) were given without characters and tones.

<sup>23</sup>My Mandarin informant made me aware of the fact that 狒 [*fei*] is usually employed as a character for 'to say' in Gansu province.

5 Survey of the grammars of questions in Northeast Asia

Table 5.98: Hezhou/Linxia Chinese question markers (Xie Xiaolan & Zhang Shumin 1990: passim); notation of vowels slightly adjusted; elements given in characters only are rendered here in Pinyin without tones in square brackets as an approximation

Type	Form	Usage
PQ	<i>ma</i> <sup>3</sup>	with negation
	<i>la</i> <sup>3</sup>	with assertion
	<i>ni</i> <sup>3</sup> <i>mu</i> <sup>3</sup>	polite, solemn, younger towards older speakers
CQ	<i>za</i> <sup>3</sup>	PERSON, THING, QUANTITY
	<i>ni</i> <sup>3</sup>	PLACE, TIME, MANNER, male and young speakers
	<i>ni</i> <sup>3</sup> <i>za</i> <sup>3</sup>	PLACE, TIME, MANNER, female and old speakers, more polite than <i>ni</i> <sup>3</sup>
AQ	X <i>ni</i> <sup>3</sup> <i>mu</i> <sup>3</sup> , Y <i>ni</i> <sup>3</sup>	anticipated actions
	X <i>li</i> <sup>3</sup> <i>mu</i> <sup>3</sup> , Y <i>li</i> <sup>3</sup>	past actions
	X <i>ni</i> <sup>3</sup> <i>mu</i> <sup>3</sup> , Y <i>ni</i> <sup>3</sup> , [ <i>haishi</i> ] Z	Mandarin <i>háishi</i> ‘or.?’
	X <i>ni</i> <sup>3</sup> , Y, [ <i>haishi</i> ] Z	
NAQ	X <i>ni</i> <sup>3</sup> <i>mu</i> <sup>3</sup> , NEG (X)	
	X <i>la</i> <sup>3</sup> , NEG (X)	
?TQ	<i>ni</i> <sup>3</sup> <i>sa</i> <sup>3</sup>	
	( <i>ts</i> ) <sup>3</sup> <i>ni</i> <sup>3</sup> <i>sa</i> <sup>3</sup>	
	[ <i>jiushi</i> ] <i>la</i> <sup>3</sup>	Mandarin <i>jiùshi</i> ‘exactly’
	[ <i>duizhe</i> ] <i>la</i> <sup>3</sup>	Mandarin <i>duì</i> ‘correct’
	[ <i>bushi</i> ] <i>pe</i> <sup>3</sup>	Mandarin <i>búshi</i> ‘isn’t’
	[ <i>jiushi</i> ] <i>la</i> <sup>3</sup> [ <i>shi</i> ]	relative certainty
X <i>la</i> <sup>3</sup> [ <i>bu</i> ](X) [ <i>shi</i> ]	Mandarin <i>bù</i> ‘NEG’	
	<i>sa</i> <sup>3</sup>	

(270) Hezhou/Linxia

a. *fatsi-xa ke-vẽ-li=mu?*

house-ACC build-finish-PFV=Q

‘Is the house built?’

b. *ni tham-xa fima-la khuẽ-te-li?*

2SG 2PL-ACC what-INST wait-?-?INT

‘What would you serve them with?’ (Lee-Smith 1996b: 866, 868)

Dwyer (1995: 158) claims that the **Xunhua** subdialect of the Hezhou/Linxia Chinese has a tag question marker that derives from an interrogative meaning ‘what’. The only

example given is a content question, however, and more likely, it is cognate with  $z\alpha^3$  seen in Table 5.98. Content questions may also be unmarked.

(271) Hezhou/Linxia (Xunhua)

- a.  $\eta\alpha^{24}$ - $m\bar{e}^{44}$       *phai-ṣā aazi<sup>14</sup> gə<sup>23</sup> ʔpə<sup>44</sup> dzə<sup>23</sup>-gə s[ʈʰr̩ s]*      *xew sa<sup>41</sup>?*  
 1SG-PL    send-RES which    accomplish this-CLF matter      COP good Q

‘Whom should we send to take care of this?’

- b.  $\eta i^{53}$   $\mathfrak{s}\alpha^{13}$   $ma^{41}$   $k\alpha$   $mai^{53}$   $li\alpha?$   
 2SG what      buy    PFV

‘What did you buy?’ (Dwyer 1995: 158, 162)

The negator in the following negative alternative question (A-not-A) (272) is what has been called a potential (POT) meaning, cf. Mandarin *zhǎo-bu-dào* ‘seek-NEG.POT-RES’ ‘not be able to find’. But the affirmative counterpart would usually require another marker to substitute for the negator, cf. Mandarin *zhǎo-de-dào* ‘seek-NEG.POT-RES’ ‘be able to find’ (Sun Chaofen 2006: 60f.).

(272) Hezhou/Linxia (Xunhua)

- $d\mathfrak{z}\alpha^{24}$ - $g\gamma$  ʔdzūən<sup>41</sup>-dzi xē<sup>42</sup>,  $\eta i^{53}$   $na^{14}$ - $x\alpha^{41}$ - $la$   $na^{14}$ - $bu$ - $x\alpha^{41}$ ?  
 this-CLF heavy-EXT    very 2SG carry-RES-Q    carry-NEG.POT-RES

‘This is very heavy, can you carry it or not?’ (Dwyer 1995: 173)

Following Xie Xiaolan & Zhang Shumin (1990), the marker *-la* has been reanalyzed as a question marker here. Dwyer (1995) does not give an example of a polar question.

**Wutun** has a polar question marker *-a* that has been compared to both Mandarin *吗* as well as the interrogative *a-* ‘which’ (Mandarin *nǎ* 哪) (Janhunen et al. 2008: 99). Another possible source might be Mandarin *=(y)a* 啊/呀. The marker has been called an enclitic particle, but was written attached to the preceding word with a hyphen. It is analyzed as enclitic here and thus written as *=a*. Similar to some Mongolic languages in the area, the marker fuses with certain preceding suffixes, which speaks in favor of an analysis as a suffix. For instance, the continuative marker *-zhe*, combined with the question marker results in the form *-zha*, which is reminiscent of Mongolic languages of the region (§5.8.2). Apart from *=a*, Wutun allegedly has borrowed the suffix *-mu* from a Mongolic source, probably Bonan (Sandman 2012: 384). But a Turkic origin of the Wutun form is more likely, e.g. Salar *=mu*, Uyghur *=mu* (§5.11). Consequently, it has been reanalyzed as enclitic here. According to Lee-Smith & Wurm (1996: 894) it has the form *-mə* and is cognate with Mandarin *ma*, which seems unlikely but not impossible. As seen above, it has the form  $m\alpha^{24}$  in Xining Mandarin. Wutun lacks a tonal distinction.

(273) Wutun

- a. *je ni-de huaiqi hai-li=a?*  
 this 2SG-GEN book EQU-SEN-INF=Q

‘Is this your book?’

5 Survey of the grammars of questions in Northeast Asia

b. *sama qe-lio=mu?*

food eat-PFV=Q

‘Have you eaten the food?’ (Janhunen et al. 2008): 99, 100)

The functional distribution of the two suffixes remains somewhat unclear, but =*mu* is said to have been encountered less frequently (Janhunen et al. 2008). Erika Sandman (p.c. 2016) informed me that =*a* is used with imperfective as well as progressive, and =*mu* with perfective as well as resultative aspect. Content questions remain unmarked.

(274) Wutun

*ni ma-ge nian-di-yek?*

2SG what-CLF read-PROGR-EGO

‘What are you reading?’ (Janhunen et al. 2008: 98)

The following examples in (275) were kindly provided by Erika Sandman (p.c. 2016), see also Sandman (2016: 287–297).

(275) Wutun

a. *ni xaitangwa hai-yek=mu ggaigan hai-yek?*

2SG student EQU-EGO=Q teacher EQU-EGO

‘Are you a student or a teacher?’

b. *zang jja-la-gu-la waiqai yek, mi-yek?*

PN visit-INCOMPL-COMP-COND hardship EX NEG.EX-?EGO

‘If you visited Tibet, would there be any hardships?’

c. *ni yan za-de yek ya?*

2SG tobacco smoke-N EX Q

‘You are a smoker, aren’t you?’

d. *nia-ha dun-li=a?*

2SG.OBL-TOP cold-SEN.INF=Q

‘And how about you, are you feeling cold?’

Whether this last example (275d) can be analyzed as focus questions remains somewhat dubious. Similar to Japanese (§5.6.2) and Korean (§5.7.2), it is perhaps better analyzed as a polar question with an additional topic. Similar to Tibetic there seems to be interaction with egophoricity (§5.9.2.2). The egophoric marker *-yek* (< *yek* = Mandarin *yǒu* 有) is “typically used with the first person in statements when the action is volitional and allows the speaker’s control” (Sandman 2016: 209). There are certain other reasons in which the egophoric marker can be used for non-first person (Sandman 2016: 222), but none of them seems to apply in examples (274), (275a), or (276).

(276) Wutun

- a. *ngu-de minze-li dongzhek sho-yek.*  
 1SG-GEN name-LOC PN say-EGO  
 ‘My name is Dongzhou.’
- b. *ni-de minze-li ma sho-yek.*  
 2SG-GEN name-LOC what say-EGO  
 ‘What is your name?’ (Sandman 2016: 294f.)

Thus, Wutun appears to follow the anticipation rule as described in §4.4 (Tournadre & LaPolla 2014: 245; Sandman 2016: 294). Sandman (2016: 226) mentions another interesting interaction with evidentiality: “The questions with factual evidential [*re*] differ from questions that are true requests for information. In my data, factual evidential was used in questions with an obvious answer”.

**Tangwang** has also adopted the Uyghur polar question marker =*mu* and has unmarked content questions. Examples provided by Xu (2014) were given without tones. Lee-Smith (1996c) does not mention any questions.

(277) Tangwang

- a. *ni nə tsuə=mu?*  
 2SG be.able ?live=Q  
 ‘Can you live?’ (Xu 2014: 357)
- b. *ni-m tcia.tsu tci xuə jəu-li, mə-jəu?*  
 2SG-PL family GEN life EX-?IPFV NEG-?EX  
 ‘Does your family have a life (worth living) or not?’
- c. *tʂ<sup>31</sup>-fiə<sup>31</sup> ʂu<sup>24</sup>-xa a<sup>24</sup>m mə<sup>31</sup>-li?*  
 this-PL book-?TOP how sell-IPFV  
 ‘How does one sell these books?’ (Yibulaheimai A. 1985: 36)

There are also very few recordings of questions in **Gangou** Chinese. Content questions remain unmarked. Alternative questions appear to have just one marker on the first alternative. Polar questions most likely have the sentence-final marker *ma*, but no example was found.

(278) Gangou

- a. *aijie amen mei-lai shuo?*  
 3SG how NEG-come HS  
 ‘Why didn’t he come?’
- b. *ni liang.ge huaer yao-li ma, bu-yao?*  
 2SG two flower want-FUT Q neg-want  
 ‘Do you two want a flower or not?’ (Zhu Yongzhong et al. 1997: 447, 437)

5.9.2.2 Question marking in Tibetic and Qiangic

Ebihara (2011: 64f.) mentions several different ways of forming questions in **Amdo Tibetan**. Polar and content questions may optionally take a sentence-final clitic =*ni*. Its origin eludes me, but one might compare it with Hezhou/Linxia *ni*<sup>3</sup> 呢.

(279) Amdo Tibetan (Gonghe; Tibetic)

a. *te<sup>h</sup>o wə joŋ=ni?*  
2SG go.out.PRF come=Q

‘You came back?’

b. *te<sup>h</sup>əzəka ma-joŋ=ni?*  
why NEG-come=Q

‘Why did (you) not come?’ (Ebihara 2011: 65)

Content questions may also remain unmarked and polar questions have another enclitic =*na*. The distribution of the two markers among polar questions remains unclear.

(280) Amdo Tibetan (Gonghe; Tibetic)

a. *te<sup>h</sup>o s<sup>h</sup>ə jən?*  
2SG who COP.CJ

‘Who are you?’

b. *te<sup>h</sup>o demo jən=na?*  
2SG fine COP.CJ=Q

‘Are you alright?’ (a greeting) (Ebihara 2011: 65)

Polar questions alternatively may be marked with intonation exclusively. There is yet another marking strategy for polar questions that is almost unique within the Northeast Asian area: Amdo Tibetan possesses a verbal prefix for marking polar questions. Again, there is no comment on the functional distribution of this marking strategy with respect to the others. But this might be the default marking.

(281) a. Amdo Tibetan (Gonghe; Tibetic)

*te<sup>h</sup>o wol ə-jən?*  
2SG PN Q-COP.CJ

‘Are you Tibetan?’ (Ebihara 2011: 70, 64)

b. *norbə joŋ=dzi ə-re?*  
PN come=N Q-COP.DJ

‘Will Norbu come?’

(282) Amdo Tibetan (dPa’ris; Tibetic)

*te<sup>h</sup>o: χ<sup>w</sup>ie<sup>h</sup>a ə-jol?*  
2SG.DAT book Q-COP.CJ

‘Do you have a book?’ (Ebihara 2013: 155)

In the examples provided by Ebihara, the prefix always attaches to a copula, but it is not restricted to this context. Consider two examples from the Themchen dialect (283).

(283) Amdo Tibetan (Themchen)

a. *tɕ<sup>h</sup>o ts<sup>h</sup>uŋwa ɶ-jən?*  
 2SG merchant Q-COP.CJ  
 ‘Are you a merchant?’

b. *bdəstɕəl ɶ-t<sup>h</sup>a?*  
 PN Q-go.PFV.DJ  
 ‘Has Dekyi gone?’ (Haller 2004: 69, 81)

Janhunen (2012c: 184) is correct that the prefix represents an important difference of Amdo-Tibetan when compared with the other languages of the Amdo Sprachbund. But as seen before, Amdo Tibetan has sentence-final particles as well, and as we will see further below there are other languages in the region with a similar pattern. Basically, the same pattern as in the dialects mentioned above is also found in other dialects such as that of Tongren/Rebgong Amdo Tibetan, e.g. *tɕ<sup>h</sup>o demō jin=na?* ‘How are you?’ (see 280b), and *e-jol* ‘Q-COP.CJ’ (see 282) (de Roerich 1958: 98, modified transcription). The descriptions of Tibetic languages included here usually do not mention alternative, focus or tag questions. However, Tongren/Rebgong Amdo Tibetan has a tag question marker *e-den-gA* ‘Q-truth-?GEN’ (de Roerich 1958: 131, modified transcription).

Amdo Tibetan has a distinction between conjunct and disjunct marking that usually is manifested in the copula system. The distinction has also been adopted from Tibetic by several Mongolic (§5.8.2) and Sinitic (see above) languages of the Amdo Sprachbund (§3.5). According to Aikhenvald (2012: 471) “the alternation between conjunct and disjunct person marking marks new information and surprise, especially in 1st person contexts. The disjunct person marking indicates something out of the speaker’s control, unexpected and thus surprising.” As we have already seen in §4.4, there is some interaction of conjunct/disjunct marking and questions: “In question sentences for the second person, the conjunct forms are generally used according to the point-of-view of the second person” (Ebihara 2011: 69). In Gonghe Amdo Tibetan, for example, there are special conjunct (*jən*, NEG *mən*) and disjunct (*re(l)*, NEG *mare(l)*) copula forms (Ebihara 2011: 69), see also examples (280,281, 282,283a) above.

The intonation of Amdo Tibetan questions has been given in quite some detail by Sun (1986) for the dialect spoken in Xəra village in Northern Sichuan:

The interrogative word is spoken on a high falling pitch, or, if the interrogative word has more than one syllable, a high falling pitch on the last syllable and high level pitch on the other syllables. [...] The typical intonation of yes-no questions is a high level pitch on /s/ followed by a high falling tune realized on the verbal element. (Sun 1986: 60f.)

What is given here as /s/ corresponds here to the question prefix *ɶ-* in Gonghe and other dialects (cf. Sun 1993: 959). According to Denwood (1999: 128), Lhasa Tibetan <e> /ʔʌ/,

apparently a cognate of *a-*, generally has dubitative meaning but also functions as a polite question marker for the second person.

Classical Tibetan has a sentence-final question marker *'am* that also assimilates to the preceding word. Content questions remain unmarked.

(284) Tibetan (Classical; Tibetic)

a. *mi de dgra yin=nam?*  
 person that enemy COP=Q

'Is that man an enemy?'

b. *kyod kyis glang brnyas=sam?*  
 2SG ERG OX borrow=Q

'Did you borrow (his) ox?'

c. *nga-s khyod kyi khyo ci.ltar sbyin?*  
 1SG-ERG 2SG GEN husband how give

'How can I give you back your husband?' (DeLancey 2003: 262, 267)

According to DeLancey (2003: 267), the Classical Tibetan sentence-final polar question marker *'am* "represents a reduction of an earlier balanced question construction, probably \*V 'o ma-V 'V(or) not-V?' > V 'am 'V". Thus, the development is from a negative alternative question construction to a polar question marker. Amdo *=na* seems to correspond to Classical Tibetan *=nam* (de Roerich 1958: 98).

Zhongu (286), Baima (287), and Tangut (288) also have a verbal prefix for polar questions and unmarked content questions. But Tangut, like Amdo Tibetan, also has a sentence-final question marker *mo*<sup>2</sup>, which formally looks similar to the marker in Xining Mandarin (284).

(285) Tangut (Qiangic)

*nji<sup>2</sup> kã<sup>1</sup> tšja<sup>1</sup> wji<sup>2</sup> dzjo<sup>1</sup> mo<sup>2</sup>?*  
 2SG.HON sugar cane PERF eat Q

'Do you eat sugar cane?' (Gong Hwang-Cherng 2003: 614)

The Baima question marker has been reanalyzed as a prefix here in analogy to the other languages.

(286) Zhongu (?Tibetic)

a. *ɸ-sə-kə?*  
 Q-sour-MIR

'Is it sour?'

b. *ts<sup>h</sup>o(-sɸ) gomo tʃ<sup>n</sup>atsə <sup>n</sup>dərə?*  
 2SG(-DAT) money how.much EX

'How much money do you have?' (Sun 2003b: 826)

- (287) Baima (?Tibetic)
- a. *tcho*<sup>13</sup> *ko*<sup>53</sup> *ŋge*<sup>13</sup> *re*<sup>35</sup> *e*<sup>53</sup> *-ndzi*<sup>53</sup> *i*<sup>53</sup>?
- 2PL PN Q-go ?PROGR
- ‘Are you going to Baima?’
- b. *tchø*<sup>53</sup> *su*<sup>341</sup> *re*<sup>13</sup>?
- 2SG who COP
- ‘Who are you?’ (Sun Hongkai et al. 1996: 131, 126)
- (288) a. Tangut (Qiangic)
- mə*<sup>1</sup> *ʃij*<sup>1</sup> *ɣu*<sup>1</sup> *ʃa-tchji*<sup>1</sup> *-dju*<sup>1</sup>?
- sky GEN head Q-?POT-have
- ‘Does the sky have a head?’ (Jacques 2011: 427)
- b. *thji*<sup>2</sup> *ʃjo*<sup>2</sup> *thji*<sup>2</sup> *dzjwo*<sup>2</sup> *tji*<sup>1</sup> *ʃij*<sup>1</sup> *rjur*<sup>1</sup> *ar*<sup>2</sup> *mji*<sup>1</sup> *njwi*<sup>2</sup> *ʃj*<sup>2</sup>?
- why this person alone ACC restrain NEG can COMP
- ‘Why can’t you restrain this person alone?’ (Gong Hwang-Cherng 2003: 614)

It seems possible that the preverbal question marker is an areal feature. Some other Qiangic languages share the same question marking strategy as well. Since all Qiangic languages today are located outside of NEA, some examples should suffice (289, 290, 291).

- (289) Caodeng rGyalrong (Qiangic)
- neji*? *ne-lŋa*? *ʃ-to*??
- 2SG 2SG.POSS-child Q-exist
- ‘Do you have children?’ (Sun 2003a: 498)
- (290) Jiaomuzu rGyalrong (Qiangic)
- no tʃha wutə ʃ-tə-mut-w*?
- 2SG tea this Q-2-drink-2SG
- ‘Will you drink this tea?’ (Prins 2017: 593)
- (291) Guiqiong (Qiangic)
- zo gutchiŋ ɛ-dzi*?
- 3SG PN Q-COP
- ‘Is he a Guiqiong?’ (Jiang Li 2015: 304)

It can also be found in further Amdo Tibetan dialects and other Tibetic languages of the region such as the gSerpa variety in northwestern Sichuan.

- (292) Amdo Tibetan (Xəra; Tibetic)
- tç<sup>h</sup>o xabda ʃ-s<sup>h</sup>oŋ*?
- 2SG deer.chase Q-go.COM
- ‘Did you go deer-hunting?’ (Sun 1993: 959)

5 Survey of the grammars of questions in Northeast Asia

(293) gSerpa Tibetan (Tibetic)

*martsə hətʰo-lə rdzəvuuə χtsə-ve ə-γəʔ*

originally 2PL.family-DAT louse one-except Q-COP

‘So there was just one single ‘louse’ in your home?’ (Sun 2006: 125)

Possibly, there are areal connections to Chinese dialects, too. Consider the following example from the Hefei dialect spoken in central Anhui province.

(294) Chinese (Hefei)

[ni *ke-xiang.xin*]

2SG Q-believe

‘Do you believe (that)?’ (Dexi 1985: 12)

In this dialect the marker has the form *kʰəʔ¹* or *kəʔ¹*. An investigation of the extent of this feature towards the south goes beyond the possibilities of this study. But at least Mandarin as spoken in Yunnan also has this pattern. Independent of that question, it represents a southern border of the NEA area as no other language in the sample has a comparable pattern. It may also be noted that Qiang, the southern neighbor of Baima and Zhongu does not share this pattern (§4.2.3, LaPolla & Huang Chenglong 2003: 180). The marker sometimes can exhibit a rather complex morphosyntactic behavior. For example, in Prinmi, a language also spoken in Yunnan, the marker has the form *a* and usually, but not always, takes penultimate position in a sentence (Ding 2014: 208).

(295) Prinmi (Qiangic)

a. *pə̌ʰpə̌ʰ aʰ=ʒeʰ?*

uncle Q=EX.AN

‘Is Uncle home?’

b. *tʰeʰ dzɿʰ=aʰ=fəʰ?*

meal eat=Q=OPT

‘Will (you) have a meal?’ (Ding 2014: 209)

Consequently, the marker can stand both in front of or after the verb. In Japhug (rGyalrong, Qiangic), to mention yet another language from Sichuan with the feature in question, the prefix apparently invariably has the form *u-* (and usually receives stress) (see Jacques 2004: 400f.). However, the form of the marker is sometimes variable. In Baima the question marker has several different variants shown in Table 5.99 that are determined by the vowel of the following verb, i.e. it exhibits some form of umlaut. However, unlike Prinmi, there is no change in tone.

Whether the markers in all languages mentioned above actually are cognates of each other could not be settled here but seems likely except for Chinese. However, this is irrelevant from an areal and typological perspective.

Table 5.99: Variants of the question marker in Baima (Sun Hongkai et al. 1996: 85).

Example	Meaning
e <sup>53</sup> -ndzi <sup>53</sup>	Q-go
ɛ <sup>53</sup> -tʃe <sup>53</sup>	Q-cut(.tree)
ɛ <sup>53</sup> -nɛ <sup>35</sup>	Q-sleep
a <sup>53</sup> -dʒa <sup>341</sup>	Q-mend
a <sup>53</sup> -tɔ <sup>35</sup>	Q-wrap
ɔ <sup>53</sup> -kho <sup>53</sup>	Q-back
o <sup>53</sup> -phu <sup>35</sup>	Q-rub
ø <sup>53</sup> -ny <sup>341</sup>	Q-smell
ə <sup>53</sup> -khe <sup>53</sup>	Q-toast
ə <sup>53</sup> -ndzø <sup>53</sup>	Q-hit
ə <sup>53</sup> -dzue <sup>341</sup>	Q-pick(.potatoe)

### 5.9.2.3 Summary

Table 5.100 summarizes the marking of polar and content questions in Trans-Himalayan languages located in Northeast Asia. Clearly, there is a tendency for marked polar questions and unmarked content questions.

Table 5.100: Polar and content question markers in Trans-Himalayan languages spoken in NEA. Tones are often variable and were thus excluded here.

	PQ	CQ
Old Chinese	*çʰa [hu]#	-
Mandarin	ma#	-, ne#
Urumqi Hui Chinese	ma#	-, nə# ~ ni# ~ niɛ#
Xining Mandarin	mɔ#	-, lɛ#
Gangou	?ma#	-
Hezhou/Linxia	ma#, la#, (ni)mu#	-, za#, ni#, niza#
Wutun	PFV, RES =mu#, IPFV, PROGR =a#	-
Tangwang	=mu#	-
Amdo Tibetan	ə-v, =na#, =ni#	-, =ni#
Zhongu	v-v	-
Baima	e-v (see Table 5.99)	-
Tangut	ʃja-v, mo#	-

## 5.9.3 Interrogatives in Trans-Himalayan

## 5.9.3.1 Interrogatives in Sinitic

Sinitic, and particularly Mandarin, interrogatives are special in several regards. First, modern Mandarin interrogatives differ strongly from those in Old Chinese. Second, many of the forms are transparent formations that indicate a recent origin. Let us first address the Old Chinese interrogatives. Table 5.101 gives their recent reconstruction in the Baxter Sagart system.

Table 5.101: Interrogatives in Old Chinese according to Baxter & Sagart; Baxter & Sagart's (2014a; 2014b) reconstruction; Middle Chinese is only an approximation; square brackets indicate uncertain sounds; forms marked by a question mark were not actually reconstructed as interrogatives by Baxter & Sagart (cf. Pulleyblank 1995: 91–97)

Meaning	Character	Modern Reading	Middle Chinese	Old Chinese
how	安	ān	'an	*[ʔ] <sup>s</sup> a[n]
how	焉	yān	'jen	*ʔa[n]
what	何	hé	ha	*[g] <sup>s</sup> aj
what	曷	hé	hat	*[g] <sup>s</sup> at
(dial. for 何)				
why, what	奚	xī	hej	*[g] <sup>s</sup> e
how many	幾(几)	jǐ	kj+jX	*kəjʔ
how	豈	qǐ	khj+jX	*C.q <sup>h</sup> əjʔ
(rhetorical)				
who	孰	shú	dzyuwk	*[d]uk
who	誰	shuí	dzywij	*[d]uj
who	疇	chóu	?drjuw	?*[d]ru
why, how	胡	hú	?hu	?*[g] <sup>s</sup> a
why	胡為	húwèi	?hu + hjwe	?*[g] <sup>s</sup> a + *C <sup>w</sup> (r)aj
why not	盍	hé	?hap	?*m-[k] <sup>s</sup> ap
(< 何不?)				
how, where	惡(乎), 烏	wù(hú), wū	?uh, 'u	?*ʔaks, *[ʔ] <sup>s</sup> a
(< 於何?)				

Thus, Old Chinese may have had resonances in \*ʔ-, \*d-, and \*g<sup>s</sup>- as well as several other interrogatives without such a submorpheme (cf. Pulleyblank 1995: 91). Regardless of whether the pharyngealization hypothesis (indicated with /<sup>s</sup>/) turns out to be true (Baxter & Sagart 2014a: 68ff.), several forms qualify as *K*-interrogatives. Perhaps, the polar question marker \*C<sup>w</sup>a 乎 also belongs here. The exact analysis of most forms is unclear. But note that at least in some interrogatives analyzable morphological elements may have been present. The difference between \*[d]uk 孰 and \*[d]uj 誰 is especially

intriguing. According to Pulleyblank (1995: 92) the former belongs to “a group of words in \*-k [...], which are confined to preverbal position referring to the subject, and which usually select the subject from a larger group.” Xu (2006: 236ff.) agrees in part with this assessment and argues that a group of words ending in \*-j (such as \*[d]uj 誰) are more flexible in their syntactic behavior than those in \*-k (such as \*[d]uk 孰). Pulleyblank (1995: 91) furthermore assumes that several interrogatives including *ān* 安 and *yān* 焉 are derived from the “coverb” (perhaps better called preposition) *yú* 於 in combination with unspecified other elements. In fact, Baxter and Sagart reconstruct the form as \*[?]<sup>6</sup>a 於, which makes a connection with \*[?]<sup>6</sup>a[n] 安 and \*<sup>6</sup>a[n] 焉 seem possible. But if this assumption is true, the preposition must have fused with a following interrogative. These approaches are far from offering a clear picture of the etymology or morphology of Old Chinese interrogatives. To track the development of interrogatives—or of questions in general for that matter—goes well beyond the possibilities of this study (but see Peyraube & Wu 2005).

Colloquial **Mandarin Chinese** (Table 5.102) potentially has only one interrogative that is synchronically non-analyzable, namely *shéi* (*shuí*) ‘who’. All other interrogatives are analyzable to different degrees. Some are straightforward combinations of an interrogative and a noun such as *shénme dìfang* ‘where’. The second part simply means ‘place’, but the first element *shénme* ‘what’, like *zěnmé* ‘how’ possibly contains a suffix *-me* with an opaque meaning. In the complex interrogative *zěnmé*-*yàng* ‘how, what kind of, in what way’, the element *-me* may be omitted, which speaks in favor of an analysis as a suffix. Other interrogatives productively combine with grammatical elements such as the classifier *ge* 个. A special case is the interrogative *gàn.má* ‘to do what, why’, which quite clearly is a contraction of the transparent formation *gàn shénme* ‘to do what’. The first element *gàn* ‘to do’ remains transparent, but the second element *má* is what is usually called a cranberry morph, because it is not attested outside of this word. The interrogatives *jǐ*- ‘how many’ or *nǎ*- ‘which (one)’ do not qualify as “basic question words” either, because they necessarily combine with another element such as a classifier. The lack of a strongly developed resonance speaks in favor of a relatively new system of interrogatives. In fact, only *shéi* (*shuí*) 誰 and *jǐ*- 几 can be traced back to Old Chinese.

However, apart from the interrogatives mentioned in Table 5.102, Mandarin has about a dozen or so formal interrogatives given in Table 5.103 that are mostly restricted to the literary language and preserves Old Chinese \*<sup>6</sup>[g]<sup>6</sup>aj 何.<sup>24</sup>

To my knowledge, in NEA only Mandarin has such a marked contrast between two different sets of interrogatives that depend on style.

Mandarin interrogatives display strong paradigmatic similarities with the demonstratives. Mandarin is not usually analyzed as having paradigms, but nevertheless such an analysis seems viable (Table 5.104).

<sup>24</sup>It may be noted that Japanese also preserves the character 何 but has an autochthonous pronunciation *nani* なに instead.

5 Survey of the grammars of questions in Northeast Asia

Table 5.102: Mandarin Chinese interrogatives and their analysis (mostly based on own knowledge, elicitation, Ross & Ma 2006: 160f.); not all combinations are shown

Meaning	Character	Form	Analysis
how + ADJ	多	duō + ADJ	< ‘very’ < ‘much’
e.g., how long	多长	duō-cháng	<i>cháng</i> ‘long’
how many/much	多少	duō-shǎo	<i>duō</i> ‘much’, <i>shǎo</i> ‘few’
how + ADJ	好	hǎo- + ADJ	< ‘very’ < ‘good’, dialectal variant of <i>duō</i>
how many/much	几	jǐ- + CLF	usually <i>jǐ-ge</i> 几个
at what time	几点(钟)	jǐ-diǎn(zhōng)	<i>diǎn(zhōng)</i> ‘o’clock’
which (one)	哪	nǎ- + CLF	usually <i>nǎ-ge</i> 哪个
where	哪里/儿/边	nǎ-li/-(e)r/-biān	<i>-li/-(e)r/-biān</i>
who	谁	shéi (shuí)	
what	什么	shén-me	- <i>me</i> as in <i>zěn-me?</i> , <-n-m-> = [-mm-]
what	啥	shá	colloquial variant of <i>shénme</i>
what	嘛	mà	colloquial variant of <i>shénme</i>
where	什么地方	shénme dìfang	<i>dìfang</i> ‘place’
when	什么时候	shénme shíhou	<i>shíhou</i> ‘time’
why	为什么	wèi-shénme	<i>wèi</i> ‘for’
to do what, why	干什么	gàn shénme	<i>gàn</i> ‘to do’
to do what, why	干嘛	gàn.má	colloquial variant of <i>gàn shénme</i>
on what basis	凭什么	píng-shénme	<i>píng</i> ‘rely on’
how, why	怎么	zěnme	- <i>me</i> as in <i>shén-me?</i> , <-n-m-> = [-mm-]
how	咋	zǎ	colloquial variant of <i>zěnme</i>
how	怎(么)样	zěn(me)-yàng	<i>yàng</i> ‘kind’

Table 5.103: Formal or literary Mandarin Chinese interrogatives (Ross &amp; Ma 2006: 161f.; Pulleyblank 1995, partly elicited); not all forms listed

Meaning	Character	Form	Base	Meaning
why	何必	hé-bì	bì	necessarily
why not	何不	hé-bù	bù	not
when (rhetorical)	何曾	hé-céng	céng	once
how could (one) not	何尝	hé-cháng	cháng	once
why not	何妨	hé-fāng	fāng	impede
why	何故	hé-gù	gù	reason
when	何时	hé-shí	shí	time
what is	何为	hé-wéi	wéi	cop
how, why	何以	hé-yǐ	yǐ	with, use
how	如何	rú-hé	rú	to be like
why	为何	wèi-hé	wèi	for

Table 5.104: Partial demonstrative and interrogative paradigms in Mandarin (my knowledge)

	this	that	which
plain	zhè-	nà-	nǎ-
X-PL	zhè-xiē	nà-xiē	nǎ-xiē
X(-NUM)-CLF	zhè(-yì)-ge	nà(-yì)-ge	nǎ(-yì)-ge
id. (fused)	zhèi-ge	nèi-ge	něi-ge
X-LOC	zhè-(e)r	nà-(e)r	nǎ-(e)r
X-LOC	zhè-li	nà-li	nǎ-li
X-LOC	zhè-biān	nà-biān	nǎ-biān
DIR-X-LOC	wǎng-zhè-(e)r	wǎng-nà-(e)r	wǎng-nǎ-(e)r
ALL-X-LOC	dào-zhè-(e)r	dào-nà-(e)r	dào-nǎ-(e)r
ABL-X-LOC	cóng-zhè-(e)r	cóng-nà-(e)r	cóng-nǎ-(e)r

The interrogative *duō* is usually used with scalar adjectives such as *duō-jiǔ* 多久 ‘how long’. Strangely, as my informant tells me, some dialects such as those of Guiyang use the cognate of Mandarin *hǎo* 好 instead. Literally, *duō* 多 means ‘much’ and *hǎo* 好 ‘good’, but notably both share an emphatic meaning of ‘very’. Sentences with both *duō* and *hǎo* can have their original meaning and may be regarded as polysemous. The reading depends on the intonation as well as the context. The following elicited example has been idealized and standardized to allow a better comparison.

- (296) Mandarin  
*zhè tiáo hé duō/hǎo kuān?*  
 this CLF river how broad  
 ‘How broad is this river?’

Guiyang is located outside of NEA, but the same phenomenon can also be observed, for example, in the Shiquan dialect in Shaanxi, which has the form *xao*<sup>55</sup> 好 (e.g., *xao*<sup>55</sup> *tciu*<sup>0</sup> 好久 ‘how long’).

There are more descriptions of interrogatives in Chinese dialects than can possibly be mentioned here. However, the majority simply rely on a transliteration with characters and do not give a phonetic transcription, which makes the data problematic at best. The following gives the interrogatives from a selection of different dialects, namely Suide 绥德 (northern Shaanxi), Shiquan 石泉 (southern Shaanxi), Yanggao 阳高 (northern Shanxi), Lingshi 灵石 (eastern Shanxi), and Xining 西宁 (eastern Qinghai). In addition, the interrogatives of Hui Chinese spoken in Urumqi 乌鲁木齐 (northern Xinjiang) are given. The list is not meant to provide an exhaustive overview, but gives an impression of dialectal variation found in northern Mandarin (Table 5.105).

There is a bewildering variety of different forms and combinations of forms that is qualitatively different from most other interrogative systems observed in NEA. Often a specific function may be expressed with a wide variety of different forms. For instance, the Yanggao dialect is said to have twelve locative forms. Only a selection of forms is included here. The Suide, Yanggao, and Lingshi dialects represent the disputed Jin dialect area that is sometimes distinguished from Mandarin. An interesting feature shared by these dialects is a final glottal stop such as in the classifier 个 (Mandarin *ge*, Shiquan *go*, Xining *kə*, Urumqi Hui Chinese *kɣ*, but Suide *kuəʔ*, Yanggao *kəʔ*, and Lingshi *kəʔ*, here given without tones). Many forms that were not listed cannot be found in Standard Mandarin. For example, Mandarin cannot use the plural marker *-men* 们 (Lingshi *ʂu*<sup>44</sup> *məŋ*<sup>44</sup>, Xining *fei*<sup>2421</sup> *mā*<sup>2454</sup>, Urumqi *sei*<sup>24</sup> *məŋ*<sup>21</sup>, Yanggao *suei*<sup>312</sup> *məŋ*<sup>31</sup>) or the classifier (yi)ge 一个 ‘one CLF’ in combination with the personal interrogative *shéi* ‘who’ (Urumqi *sei*<sup>24</sup> *kɣ*<sup>53</sup> ~ *sei*<sup>24</sup> *ji*<sup>21</sup> *kɣ*<sup>21</sup>).

A special case is Lingshi *ue*<sup>44</sup> *ʂu*<sup>44</sup> 兀谁 ‘who’, which contains a demonstrative *ue*<sup>44</sup> unknown in Mandarin. Whether Xining *a*<sup>44</sup> *mə*<sup>2444</sup> ‘why, how’ is related to Mandarin *zěnme* 怎么 ‘why, how’ or *nǎ* 哪 ‘which’ remains somewhat unclear. However, it has clear parallels in Hezhou, Wutun, Tangwang, and Gangou Chinese.

Table 5.105: A selection of interrogatives from Suide (northern Shaanxi, Weiqiang & Yongliang 2013: 51, slightly corrected), Shiquan (southern Shaanxi, Shi Feng 2009: 14), Yanggao (northern Shanxi, Sun Qinglin 2015: 150, vowel transcription slightly corrected), Xining dialects (eastern Qinghai, Zhang Chengzai 1980: passim), and Urumqi Hui Chinese (Liu Liji 1989: 160f., passim)

Mean.	Suide	Yanggao	Lingshi	Shiquan	Xining	Urumqi
who	ʃuei <sup>33</sup>	suei <sup>312</sup>	ʃu <sup>44</sup>		fei <sup>24</sup>	sei <sup>24</sup> kx <sup>53</sup>
which,		na <sup>53</sup>	? iaŋ <sup>535</sup>	la <sup>55</sup> go <sup>21</sup>	a <sup>44</sup> ko <sup>213</sup>	na <sup>21</sup> kx <sup>52</sup>
who		kəŋ <sup>33</sup>	kəŋ <sup>44</sup>			
what	ʃəŋ <sup>3</sup> .ma, ʃəŋ <sup>52</sup>	sa <sup>31</sup>	səŋ <sup>53</sup>	ʃa <sup>213</sup>	sa <sup>213</sup>	ʃl <sup>24</sup> my <sup>21</sup> , sa <sup>24</sup>
which	la <sup>213</sup>	na <sup>53</sup>	?iaŋ <sup>535</sup>	la <sup>55</sup>		
where	la <sup>213</sup> (-li)	na <sup>53</sup> lɛ <sup>0</sup>		la <sup>55</sup> li <sup>0</sup>		nɛr <sup>24</sup>
where	la <sup>213</sup> .tɛr	na <sup>53</sup> tɛr <sup>31</sup>		la <sup>55</sup> taŋ <sup>21</sup>	a <sup>4421</sup> tʃa <sup>53</sup> ɛ <sup>341</sup>	
when	ʃəŋ <sup>52</sup> ·xur		səŋ <sup>53</sup> xur <sup>44</sup>			
		təŋ <sup>213</sup> ·xur	tux <sup>31</sup> xuər <sup>24</sup>			tux <sup>21</sup> xuər <sup>44</sup>
	tɕi <sup>213</sup> sɪ <sup>33</sup>	tɕi <sup>53</sup> səŋ <sup>33</sup> sa <sup>31</sup> sɪ <sup>31</sup> xɣər <sup>24</sup>		tɕi <sup>55</sup> sɿ <sup>21</sup>		
				ʃa <sup>213</sup> sɿ <sup>21</sup> (xou <sup>0</sup> )	sa <sup>213</sup> sɪ <sup>3521</sup> xu <sup>12354</sup>	
how	tsua <sup>213</sup>	tsa <sup>312</sup> , tsa <sup>24</sup>	tʃa <sup>44</sup> tɕi <sup>44</sup>		?a <sup>4421</sup> -	tʃa <sup>52</sup> -
why	vei <sup>52</sup> ʃəŋ <sup>52</sup> , vei <sup>31</sup> ʃəŋ <sup>3</sup> .ma	(iəŋ <sup>31</sup> ) vei <sup>24</sup> sa <sup>31</sup>	iŋ <sup>535</sup> uei <sup>53</sup> səŋ <sup>53</sup>	uei <sup>21</sup> ʃa <sup>213</sup>	uei <sup>213</sup> sa <sup>213</sup>	
how	təŋ <sup>213</sup>	tux <sup>31</sup>		xao <sup>55</sup>	tu <sup>44</sup>	tux <sup>21</sup>
how	təŋ <sup>213</sup>	tux <sup>31</sup>	tei <sup>535</sup>		tu <sup>34</sup> ʃɔ <sup>53</sup>	tux <sup>21</sup> ʃɔ <sup>52</sup>
much	ʃɔ <sup>213</sup>	sɔu <sup>53</sup>	sɔ <sup>212</sup>			
how	tɕi <sup>213</sup>	tɕi <sup>53</sup>	tɕi <sup>212</sup>	tɕi <sup>55</sup>		
many						

5 Survey of the grammars of questions in Northeast Asia

Table 5.106: Xunhua and Jishishan Hezhou interrogatives (Zhong Jinwen 2007: 393f.); forms in square brackets from Dwyer (1995: 156)

Meaning	Xunhua	Jishishan
who, which	a <sup>3</sup> ɿ <sup>3(3)</sup> kə <sup>55</sup> [a-zɿ <sup>24</sup> -gə ~ a-ji <sup>24</sup> -gə]	a <sup>44</sup> ɿ <sup>2</sup> kə <sup>44</sup>
where	a <sup>3</sup> ɿ <sup>42</sup>	a <sup>44</sup> ɿ <sup>44</sup> ɿ <sup>2</sup>
how	a <sup>21</sup> mu <sup>55</sup> liəu <sup>3</sup>	a <sup>44</sup> mu <sup>44</sup> liəo <sup>02</sup>
what	ʂ <sup>2</sup> ma <sup>55</sup>	ʂ <sup>1</sup> ma <sup>44</sup>
what	[sa <sup>13</sup> ]	
what	tʂə <sup>55</sup> liəu	
what (emph.)	[ʂə <sup>13</sup> ma <sup>41</sup> kə]	
how much/many	tə <sup>24</sup> ʂo <sup>55</sup>	tuo <sup>31</sup> ʂao <sup>44</sup>
why	tʂa <sup>55</sup> liəu <sup>3</sup> ʂ <sup>3</sup>	a <sup>44</sup> mu <sup>44</sup>
where	[a <sup>14</sup> -dɿ-r]	
(?Turkic -DA)		
why	[a <sup>13</sup> -men <sup>24</sup> dzə]	

For *Hezhou* dialect only a few descriptions of interrogatives are available (Table 5.106). Both varieties of Hezhou listed have lost the initial nasal in the cognate of Mandarin *nǎ* 哪 and contain some etymologically opaque derivations.

Rimsky-Korsakoff (1994: 513, 515) mentions the two **Dungan** forms *dza* 怎 and *sa* 啥. Hai Feng (2002: 76) also only lists *tɕi<sup>41</sup>ʂ<sup>24</sup>* 几时 ‘when’ and *sa<sup>44</sup>* 啥 ‘what’, but includes a notation of tones. We have already encountered all these forms in several dialects above.

Despite the fact that some **Wutun** interrogatives are cognates with Mandarin, the overall picture is quite different. There are two new resonances being built up (*a-* and *ma-*), neither of which exists in Mandarin. Wutun has only one basic interrogative word, *ma* ‘what’, which possibly is a contracted form of Mandarin *shénme* 什么 > *mà* 嘛 ‘what’. A combination of *shénme* 什么 with *ge* 个 as in Wutun *ma-ge* is impossible in Mandarin, but not in the Xunhua subdialect of Hezhou, which has *ʂə<sup>13</sup>ma<sup>41</sup>kə* 什么个. The development of the meaning of Wutun *a-ge* from ‘which one’ to ‘who’ has parallels in several Mandarin dialects. Interrogatives in *Tangwang* are relatively straightforward. Only the origin of what appears to be a locative suffix *-tha* remains unclear for now. The description does not seem to be very reliable as individual interrogatives are given in different forms throughout the book (Xu 2014). Only for some forms tones were given, which is why they have been removed altogether. Some unclear forms were left aside.

There are no good descriptions of questions in **Gangou** Chinese. Only in the last two years have there been any studies of the language in China at all. But they are all from one and the same scholar Yang Yonglong (e.g., 2014: 244f.), who does not give sufficient information on pronunciation or grammar and for the most part employs Chinese characters for transcription and dialect for translation. According to Yang Yonglong (2014), the interrogatives can have a plural marker [-*men*] 们 that is said to be pronounced /mu/.

Table 5.107: Tangwang (Xu 2014: 222, *passim*), Wutun (Janhunen et al. 2008: 69f.; Erika Sandman p.c. 2016), and Gangou interrogatives (Zhu Yongzhong et al. 1997: 447; Yang Yonglong 2014: 244f. with Mandarin transliteration in square brackets)

Meaning	Tangwang	Wutun	Gangou
who	a-ke	a-ge	[age]
where	a-li	a-li	
where	a-tha	a-ra/la	[a-biao]
when	a-xuɪ		
how, what kind of	am(tʂɛ)		
why, how	amutʂɛ	a-menzai	a-men
what kind of	amutʂɪkə		
what	ʂəma	ma(-ge)	[sha]
when	ʂəma ʂ xəu		
how much	tuəʂɔ	do	
when		do-xige	[a-hui(e)r]
how many		jhi-ge	
why		ma-shema	

### 5.9.3.2 Interrogatives in Tibetic and Qiangic

Table 5.108 and Table 5.109 give some interrogatives from seven different **Tibetic** or **Qiangic** varieties, heuristically classified into those with and without tones.

For the gSerpa dialect from northern Sichuan, Sun (2006) only mentions the interrogative  $tɕ^hə$  ‘what’. This interrogative stem, present in all varieties mentioned here, has been reconstructed as (*\*tyi >*) *\*tɕ(h)i* ‘what’ for Proto-Tibetic, showing palatalization characteristic for Tibetic (Tournadre 2014: 114). The derived form, e.g.  $tʃ^hə-tʂə$  in Zhongu, according to Sun (2003b: 831), has the underlying Written Tibetan form *ci.cig*, in which the second element seems to be the indefinite article *cig*, derived from the numeral *gcig* ‘one’ (DeLancey 2003: 263). A parallel can be found in some Mongolic languages of the Amdo Sprachbund (§5.7.3).

The plural may be formed by reduplication, which has been adopted by some Mongolic languages of the region (§5.8.3), e.g. Gonghe Amdo Tibetan  $s^hə s^hə$  ‘who (PL)’ (Ebihara 2011: 54) or Baima  $su^{35} su^{35}$  ‘who (PL)’ (Sun Hongkai et al. 1996: 78). But in Themchen Amdo, for example, there are the plural forms  $s^hə-tɕ^hu$  and  $kaŋ-tɕ^hu$ , instead. The languages share stems for ‘who’, ‘which’, ‘when’, and ‘what’, the last of which is the basis for several derivations. For instance, Gonghe Amdo Tibetan  $tɕ^hə-zek-a$  contains a dative case that has the form  $-(k)a$  following *k* (Ebihara 2011: 60). This seems to be an exact parallel to Themchen  $tɕ^hə-zəç-a$ , Zhongu  $tʃ^hə-tʂə-jə$  (Sun 2003b: 797, fn. 51), and possibly Cone  $tɕ^hə^H-zə^L-ye^L$ . There are parallel formations in some Mongolic languages of the region (§5.8.3). The Gonghe interrogative  $tɕ^hə-gi$  ‘how’ as well as its Themchen cognate  $tɕ^hə-yi/-ji$  apparently contain a purposive or causative conjunction (Ebihara 2011: 71). Given the

5 Survey of the grammars of questions in Northeast Asia

Table 5.108: Gonghe (Ebihara 2011: 64), Themchen (Haller 2004: 51, 55, 66, passim), rNgawa (Suzuki 2006: 82), and Zhongu interrogatives (Sun 2003b: passim)

Form	Amdo (Gonghe)	Amdo (Themchen)	Amdo (rNgawa)	Zhongu
who	s <sup>h</sup> ə	s <sup>h</sup> ə	s <sup>h</sup> tu	sə
which	kaŋ	kaŋ	kəŋ	kɔ-te
where	kaŋ=na	kaŋ-na, tɕ <sup>h</sup> ə-na	kəŋ-tu	kɔ-no
how many/much	tə, tɕ <sup>h</sup> ə-mo-zek	tə, tɕ <sup>h</sup> ə-mu-(-zəç)	tɕ <sup>h</sup> i-mə-zə, kəŋtɕ <sup>h</sup> ə <sup>h</sup> tə, doqritɕ <sup>h</sup> ə	tʃ <sup>h</sup> á-tsə
what	tɕ <sup>h</sup> ə-zek	tɕ <sup>h</sup> ə(-zəç)	tɕə-zə	tʃ <sup>h</sup> ə(-tsə)
why, how	tɕ <sup>h</sup> ə-zek-a	tɕ <sup>h</sup> ə-zəç-a	kəŋ-tu	tʃ <sup>h</sup> á-tsə-jə
how	tɕ <sup>h</sup> ə-gi	tɕ <sup>h</sup> ə-yi/-ji		
when	?nem	nam	nam, kər-tu	nɔ

Table 5.109: gSerpa Tibetan (Nagano 1980: passim), Cone Tibetan (Jacques 2014: passim), and Baima interrogatives (Sun Hongkai et al. 1996: 77ff. 348f., passim)  
(<sup>H/L/N</sup> = high/low/neutral tone)

Form	gSerpa	Cone	Baima
who	su <sup>H</sup>		su <sup>35</sup> ~ su <sup>341</sup>
which	gän <sup>L</sup> ndi <sup>N</sup>		ka <sup>35</sup> le <sup>53</sup>
where	go <sup>L</sup> ne <sup>H</sup>	kaa <sup>L</sup> -nə <sup>H</sup>	ka <sup>13</sup> la <sup>53</sup>
how many/much	?çi <sup>H</sup> ra <sup>?</sup> çi <sup>N</sup> re <sup>?</sup>	tɕ <sup>h</sup> ən <sup>L</sup> -dze <sup>H</sup>	tʃh <sup>53</sup> zɔ <sup>35</sup> , ka <sup>13</sup> kɔ <sup>35</sup> , tʃh <sup>13</sup> zɔ <sup>35</sup>
what	çi <sup>H</sup> çi <sup>H</sup>		tʃh <sup>53</sup>
how	çi <sup>H</sup> ru <sup>L</sup>	tɕ <sup>h</sup> ə <sup>H</sup> -zə <sup>L</sup> -ye <sup>L</sup>	ka <sup>13</sup> tʃɔ <sup>53</sup>
why, how	çi <sup>H</sup> çi <sup>H</sup> gi <sup>N</sup> ton <sup>H</sup> da <sup>?</sup> <sup>L</sup> , čha <sup>H</sup> yin <sup>N</sup> nan <sup>H</sup>	tɕ <sup>h</sup> ɔ <sup>H</sup> -χtsə <sup>L</sup>	
when	nan <sup>L</sup>	næ <sup>L</sup> wö <sup>?</sup> <sup>L</sup> ndzə <sup>L</sup> ye <sup>L</sup>	nɔ <sup>35</sup> ndza <sup>53</sup>

analyzability of many of the forms, for example those starting with  $tc^h$ - in Amdo Tibetan, there is no clear resonance phenomenon in any of the languages mentioned. Apart from the stem *kaŋ* ‘where’, Ebihara (2009: 165) also mentions the form *kaŋ-ŋa*, in which the dative takes the regular form *-ŋa* following *ŋ* (Ebihara 2011: 60). Ebihara (2013: 157) also noted an interesting difference in forms meaning ‘whence’ with either the ablative or the genitive in different dialects of Amdo Tibetan.

(297) Amdo Tibetan (rGya ya)  
 $c^h$ o *kaŋ=ne* *joŋ=ne?*  
 2SG where=ABL come=AUX  
 ‘Where are you from?’

(298) Amdo Tibetan (dPa’ris)  
 $tc^h$ o *kaŋ=ngə* *bi=le?*  
 2SG where=GEN come=AUX  
 ‘Where are you from?’ (Ebihara 2013: 157)

There are parallel interrogative and demonstrative paradigms with a three way contrast similar to Japonic and Koreanic. Table 5.110 illustrates these with data from the Themchen dialect. Only the distal demonstrative shows an exact parallel.

Table 5.110: Demonstrative and interrogative paradigms in Themchen Amdo Tibetan (Haller 2004: 51, 64, 66)

	PROX (speaker)	PROX (hearer)	DIST	which
DAT	<i>ndə</i>	<i>tə</i>	<i>kan-a</i>	<i>kaŋ-a</i>
LOC	<i>ndə-na</i>	<i>tə-na</i>	<i>kan-na</i>	<i>kaŋ-na</i>
ABL	<i>ndə-ni</i>	<i>tə-ni</i>	<i>kan-ni</i>	<i>kaŋ-ni</i>

Finally, let us have a brief look at the interrogatives from the extinct language **Tangut**. Gong Hwang-Cherng (2003: 617, passim) mentions *sjwi*<sup>1</sup>, *sjwi*<sup>2</sup> ‘who’, *ljə*<sup>2</sup> ‘where’, *lji*<sup>1</sup> ‘which’, *wa*<sup>2</sup> ‘what’, *wa*<sup>2</sup>*zjij*<sup>1</sup> ‘how many/much’, *zjij*<sup>1</sup>-*mə*<sup>2</sup> ‘how many kinds’, and *thjij*<sup>2</sup> (*sjə*<sup>2</sup>) ‘why, how’ (1 = level tone, 2 = rising tone). The forms are very different from Tibetic and even Qiang (LaPolla & Huang Chenglong 2003: 53), which indicates a relatively long time of separation (see also Chirkova 2012). However, the Qiang interrogative system is relatively innovative with many forms, as in the Tibetic varieties above, being based on *ni(y)i* ‘what’. The Qiangic personal interrogatives seem to be among the most conservative (e.g., Qiang *sə*, Guiqiong *su* etc.) and are probably cognates of the Tangut and Tibetic forms above.

## 5.10 Tungusic

### 5.10.1 Classification of Tungusic

Recently, Janhunen (2012d: 16) suggested the following classification of Tungusic that is based mostly on previous work done by other researchers (Ikegami 1974; Lie 1978; Doerfer 1978a; Georg 2004). The designations of languages has been slightly adapted. The primary split in Tungusic is between northern and southern Tungusic, both of which subsequently went through a secondary split. Thus, there are four main branches that, following Janhunen, may be called Ewenic, Udegheic, Nanaic, and Jurchenic. There are, however, several minor problems with Janhunen's classification. For example, it does not show the strong dialectal division of some of the languages. Even and Evenki, for instance, are said to have about 12 and 50 dialects, respectively (e.g., Malchukov 1995; Atkine 1997). The classification of Solon into three different languages is too detailed, whereas the dialects of Oroqen are not even mentioned (e.g., Whaley & Li 2000, see Figure 5.5).

While the dialectal division of Nanaic, Udegheic, and Ewenic is rather well understood, there is almost no attempt at a classification of Jurchenic. The Jurchenic branch has been named after Jurchen, the oldest attested Tungusic language. Several scholars have tried to give an adequate account of the relation of Jurchen and Manchu, the second oldest attested Tungusic language. Janhunen (2012d: 6) claims that, despite "slight variation in the dialectal basis", the three Jurchenic languages Jurchen, Manchu, and Sibe "may be classified as a diachronic sequence of a single language". However, even if we consider Jurchen an archaic form of Manchu as does Janhunen, apparently following Doerfer (1978a: 12), this is imprecise and somewhat misleading. Doerfer's classification, of course, was written before the bulk of information necessary became available during the 1980s, when mainly Chinese linguists started to produce grammatical descriptions of Jurchenic varieties. Except for Sibe, these have mostly been neglected in western descriptions. I tentatively propose a new classification of Jurchenic (e.g., Hölzl 2017b; 2018a, Figure 5.6).

The exact branching structure, especially the precise relation of the three hypothetical branches, has yet to be investigated. Until recently, Alchuka and Bala were almost unknown in the West (e.g., Mu Yejun 1985; 1986; 1987; Ikegami 1999 [1993]; Hölzl 2014a: 212; Hölzl 2015a: 136 fn. 27; Hölzl 2017b; 2018a,a). Bala is basically Jurchenic but exhibits some influence from several other Tungusic languages (Mu Yejun 1985; 1986). Alchuka preserves some archaic features (e.g., an initial *k*-, and a verbal suffix *-z̄i* < \**-si*, Hölzl 2017b; 2018a), but also has unique innovations (such as the loss of several word internal consonants) and displays some interference from Manchuic. The existence of two distinct Jurchen languages has also been recognized by Kiyose (2000). They have been called Jurchen A (Bureau of Translators, Kiyose 1977) and Jurchen B (Bureau of Interpreters, Kane 1989) in analogy with similar cases, such as Tocharian A and B. Given that Sibe, located in Dzungaria since 1764, has been relatively isolated for over two hundred years and was strongly influenced by Khorchin Mongolian before that, it has to be kept apart from those dialects still located in Manchuria (e.g., Aihui, Lalin/Jing, Sanjiazi, Yibuqi). All modern varieties, except Bala and Alchuka, may be classified as Manchu di-

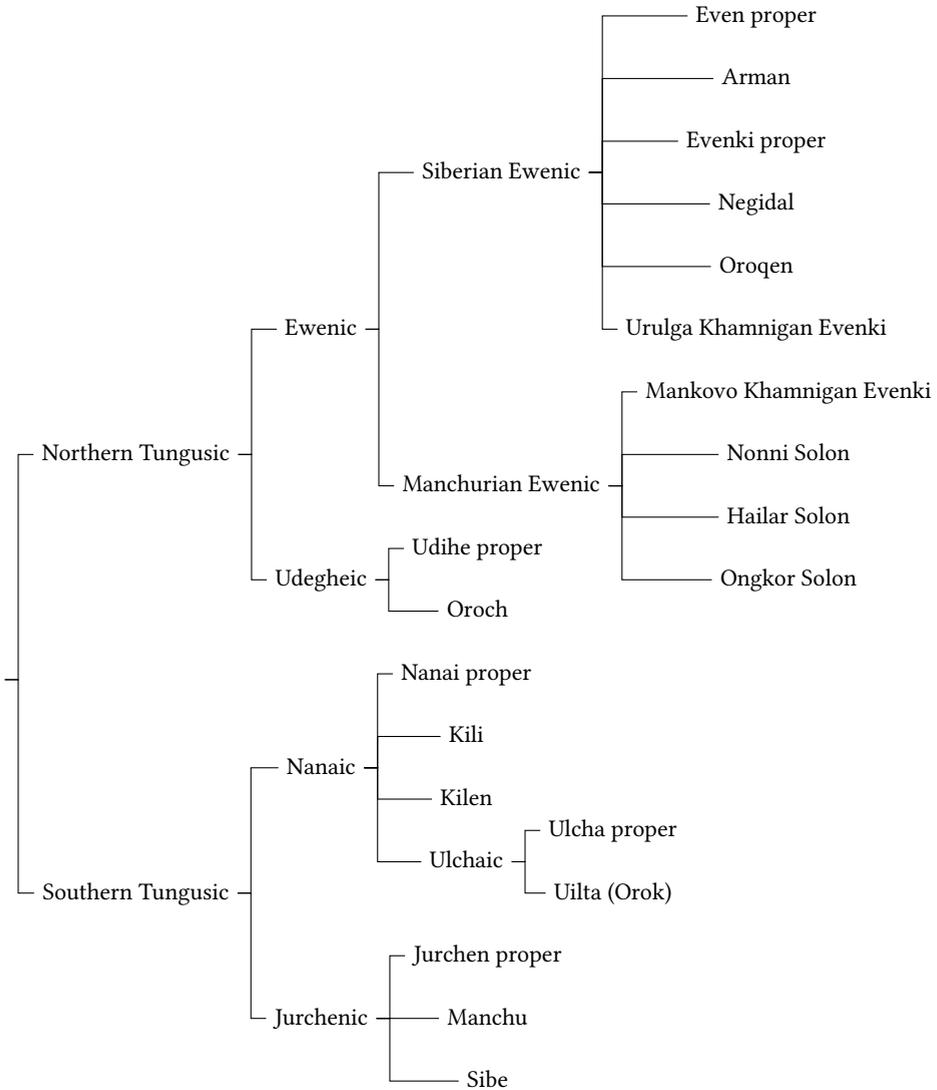
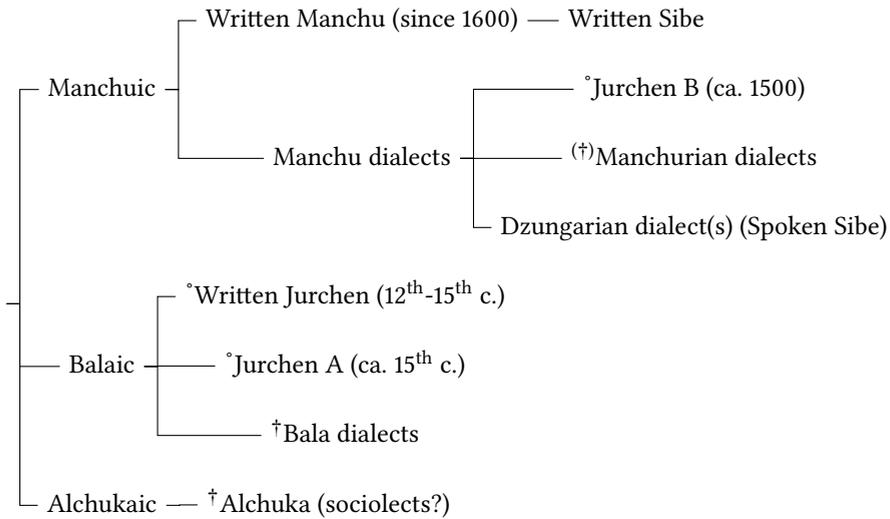


Figure 5.5: Classification of Tungusic



° = only historically attested † = no speaker left (†) = moribund, almost extinct

Figure 5.6: Proposed new classification of Jurchenic

ialects that, together with Written Manchu and Jurchen B, form the Manchuic branch of Jurchenic. Bala, together with Jurchen A (Mu Yejun 1987 also saw this connection), form a branch on their own called Balaic. The distinction between Jurchen A and Written Jurchen is mostly heuristic in nature. Technically speaking, if the above classification is correct, the forerunner of Alchuka might be called “Jurchen C” but does not seem to be attested. Only the somewhat mysterious language of the Kyakala in China (*kiyakara* in Manchu) had to be excluded for lack of data, but it seems to be a mixture of different Jurchenic varieties as well as, perhaps, some other Tungusic languages (see Hölzl 2018b for details).

### 5.10.2 Question marking in Tungusic

In *Evenki*, there are two ways of expressing polar questions. The first relies on a change of intonation: “The focus, as a rule, attracts the intonational nucleus on to itself, the intonational contour being higher and more prolonged than that of the corresponding positive sentence” (Nedjalkov 1997: 4f.). The following example can mean both ‘They killed the elk.’ (*ty* being the intonational nucleus) and ‘Did they kill the elk?’ (with the tone peak on *tyva vaa*).

- (299) *Evenki*  
*nunartyn moty-va vaa-re-Ø.*  
 3PL elk-ACC kill-NON.FUT-(3PL)  
 ‘They killed the elk./Did they kill the elk?’ (Nedjalkov 1997: 4f.)

The second way of expressing a polar question makes use of an enclitic =*Ku* that can have several variants depending on the preceding sounds, =*gu*, =*ku*, =*ŋu*, and =*vu*. The enclitic attaches to the verb in polar questions, to the focused element in focus questions, and occurs twice in alternative questions.

(300) Evenki

- a. *si ulle-ve d'ep-che-s=ku?*  
2SG meat-ACC eat-PST-2SG=Q  
'Have you *eaten* the meat?'
- b. *si ulle-ve=gu d'ep-che-s?*  
2SG meat-ACC=Q eat-PST-2SG  
'Did you eat *meat/the meat*?'
- c. *tar asatkan soŋo-d'oro-n=ŋu, in'ekte-d'ere-n=ŋu?*  
that girl cry-PRS-3SG=Q laugh-PRS-3SG=Q  
'Is that girl crying or laughing?' (Nedjalkov 1997: 136, 7)

This question marker can be traced back to Proto-Tungusic (Benzing 1956: 147) and exhibits formal and functional similarities to the Mongolic question marker (§5.8.2), which might indicate an old loan relationship of unclear direction. Buryat =*gü* and Khamnigan Mongol =*gv* might be relatively recent loans from Evenki.

Since Tungusic has different negators depending on the clause type and other factors (Hözl 2015a), negative alternative questions show different patterns as well. For standard negation, many Tungusic languages employ a negative verb. The question marker that is found once in polar but twice in alternative questions attaches to the first alternative and to the conjugated negative verb while the rest of the second alternative, including the lexical verb, is deleted. Below we will encounter negative alternative questions with other negators.

(301) Evenki

- eme-d'e-n=ŋu e.le ta.r asi, e-te-n=ŋu?*  
come-FUT-3SG-Q here that woman NEG-FUT-3SG-Q  
'I wonder if that woman will come here or not.' (Nedjalkov 1997: 7)

Content questions do not show the enclitic and remain unmarked.

(302) Evenki

- ekun-duk eme-che-s?*  
what-ABL come-PST-2SG  
'Where did you come from?' (Nedjalkov 1997: 3)

Focus questions may also remain unmarked morphosyntactically, in which case the focused element seems to take second position (cf. Nedjalkov 1997: 135).

(303) Evenki

- a. *si [er dukuvun-ma] ga-cha-s?*  
 2SG this book-ACC buy-PST-2SG  
 ‘Did you buy this book?’
- b. *[er dukuvun-ma] si ga-cha-s?*  
 this book-ACC 2SG buy-PST-2SG  
 ‘Did you buy this book?’ (Nedjalkov 1997: 5)

Question marking in Evenki dialects does not appear to differ much from Standard Evenki. The following examples were drawn from the **Sakhalin** dialect that belongs to the eastern group of dialects (Atnkine 1997). In this dialect the enclitic has a long vowel but displays the same semantic scope and morphosyntactic behavior.

(304) Evenki (Sakhalin)

- a. *kilivliil, saa-ra-s ee-ra-s=**kuu**?*  
 girl know-AOR-2PL NEG-AOR-2PL=Q  
 ‘Girls, have you noticed?’
- b. *ile-ve=**guu** e-chee-s xenu-re?*  
 man-ACC=Q NEG-PST-2SG notice-CN  
 ‘Haven’t you noticed the man?’
- c. *ηene-ηeet-y-vun=**ηuu** e-ηeet-y-vun=**ηuu**?*  
 go-DEB-E-1PL.EX=Q NEG-DEB-E-1PL.EX=Q  
 ‘Should we go or not?’
- d. ***eedaa** ekeendeek-lee-vun eme-nni?*  
 why dancing.place-LOC-1PL.EX.POSS come-2SG  
 ‘Why have you come to our dancing place?’ (Bulatova & Cotrozzi 2004: 64, 72, 19, 12)

There is also an enclitic =*too* of unclear origin that marks polar questions and does not seem to exist in Standard Evenki (Nedjalkov 1997).

(305) Evenki (Sakhalin)

- ta.r-gachiin beje-ηelii kuxii-ηeet-y-c=**too**?*  
 that-EVAL man-COM fight-DEB-E-2SG=Q  
 ‘You are supposed to fight with such a man?’ (Bulatova & Cotrozzi 2004: 19)

**Khamnigan Evenki** preserves the original enclitic as =*gv* but differs from Evenki in having borrowed both Russian =*li* (§5.5.2.2), as well as the corrogative marker *bei* from Khamnigan Mongol (§5.8.2). In Mongolic the marker is derived from the copula and even in Khamnigan Evenki seems to be mutually exclusive with the autochthonous copula *bi*. The fact that Khamnigan Evenki =*gv* does not show a variation, as in Evenki, may indicate influence from Khamnigan Mongol.

(306) Evenki (Khamnigan)

a. *aya bi-si-ndi=gv?*  
good be-PRS-2SG=Q

‘Are you well?’

b. *e.r-si nii bei?*  
this-2SG.POSS who Q

‘Who is this (one of yours)?’

c. *sii nii bi-si-ndi?*  
2SG who be-PRS-2SG

‘Who are you?’

d. *murin=li, vkvr=li?*  
horse=Q cow=Q

‘(Is it) a horse or a cow?’ (Janhunen 1991: 95f.)

In **Even** the enclitic has the variants =*gu*, =*ku*, and =*ŋu* (Malchukov 1995: 19) and as in Evenki marks polar, focus, and alternative questions. Content questions remain unmarked.

(307) Even

a. *i-d'i-m=gu?*

enter-FUT-1SG=Q

‘Shall I come in?’ (Malchukov 2001: 165)

b. *tiniv=gu em-ri-n?*  
yesterday=Q come-PST-3SG

‘Did he come *yesterday*?’ (Andrej Malchukov p.c. 2013)

c. *uliki-w=gu bu-ri-s, hulica-m=gu?*  
squirrel-ACC=Q give-PST-2SG fox-ACC=Q

‘Did you give (him/her) a squirrel or a fox?’ (Benzing 1955: 111)

(308) Even

*etiken i-le hör-re-n?*

old.man which-ALL go-NON.FUT-3SG

‘Where has the old man gone?’ (Malchukov 1995: 19)

Even has a further question marker =*i* (Malchukov 2008: 138) with possible parallels in Negidal, Solon, and, less likely, Uilta. Dialects of Even show basically the same question marking patterns. Consider the following examples from the western dialect area.

(309) Even (Western)

a. *ta.wa.r<sup>i</sup>ε.k bi-d'i-n?*

that what COP-FUT-3SG

‘What is it?’ (said in riddles)

- b. *oldo-w oldo-mi-s=gu?*  
 fish-ACC fish-v-2SG=Q

‘Do you catch fish?’ (Sotavalta 1978: 30, 28, simplified)

For the eastern dialect area (from the river Anadyr), Schiefner (1874: 217) has an example of an alternative question without a question marker but with what appears to be a disjunction *tömi*. Given that disjunctions are very rare in the northern part of NEA (§6.4), but also exist in Kolyma Yukaghir (§5.14.2), an areal connection seems possible. Malchukov (2001: 179) argues that imperative sentences in Even “may be used in interrogative sentences to ask for permission”.

- (310) Even  
*kosci-de-ku?*  
 fetch-FUT.IMP-1SG

‘Shall I fetch (the reindeer)?’ (Malchukov 2001: 165)

This might indicate a certain connection to the Chukotko-Kamchatkan languages in which there is a general affinity of imperatives to question marking (§5.3.2).

Matić (2016) claims that Even has a special category of tag questions that developed out of the negative verb.

- (311) Even (Tompo)  
*adzit=ta, ta-la ih-ha-p e-he-p?*  
 truth=?and that-ALL arrive-NFUT-1PL NEG-NFUT-1PL

‘And indeed, we have arrived there, haven’t we?’ (Matić 2016: 171)

It may be noted that the construction actually has the form of an elliptical negative alternative question (*or not?*), but with juxtaposition instead of double marking. This certainly explains the fact that, as in many other examples from Tungusic languages, the negative verb takes the same suffixes as the lexical verb. An interesting phenomenon is the optional presence of a contrastive or adversative enclitic  $C=kA \sim V=kkA$  that precedes the negative verb, but is not restricted to questions (Benzing 1955: 112).

- (312) Even (Tompo)  
*hii=kke e-he-ndi e.re.k kuŋaa-w čorda-ndi?*  
 2SG=CONTR NEG-NFUT-2SG this child-ACC beat-2SG

‘You beat up this child, didn’t you?’ (Matić 2016: 172)

The word order in this last example (312) is indeed problematic for the analysis as alternative question and strongly speaks in favor of Matić’s (2016) analysis, although there are other examples with relatively free word order above (e.g., 307).

For Arman, Doerfer & Knüppel (2013)—the only source readily available—do not have examples for any question type. Given its very close relation to Even, we may speculate that the marking of questions was similar. However, several interrogatives are attested and will be presented in §5.10.3.

Questions in Oroqen are usually marked with an enclitic that has the variants =*ηee* ~ =*ηEE* after nasals and =*jee* ~ =*jEE* in all other positions (Hu Zengyi 2001: 157). It cannot be cognate with Evenki =*Ku* which exists in Oroqen as well. Most likely it has a connection to =*yee* in the Mongolic language Dagur (but see Whaley 2005). The enclitic marks polar and alternative questions. No instance has been found where it marks focus questions.

(313) Oroqen (Nanmu)

a. *tari naan nammu-ŋi bəi=ηee?*

3SG also PN-GEN man=Q

'Is (s)he also from Nanmu?'

b. *miin əgdi nin.ut=ηee unaadzi=ηee?*

SUPERL old boy=Q girl=Q

'Is your oldest (child) a boy or a girl?' (Chaoke D. O. 2007: 141, 152)

In the Shengli dialect, the enclitic has a variant =*Ni* after nasals. This throws some doubt on the connection with Dagur but opens up the possibility of a comparison with Solon =*gi(i)*.

(314) Oroqen (Shengli)

*ɔlɔ-jɔ pi-xi-n=Ni?*

fish-PART COP-PRS-3SG=Q

'Is there any fish?' (Han Youfeng & Meng Shuxian 1993: 307)

Content questions do not have the enclitic and are unmarked morphosyntactically as in Evenki, and Even. This appears to be a difference to Dagur, but as we will see for the Xunke dialect of Oroqen below, the enclitic optionally also marks content questions.

(315) Oroqen (Chaoyang)

*fii ikʊn dʒaaln ə.lə əmə-tʃə-j?*

2SG what reason here come-PST-2SG

'Why did you come here?' (Hu Zengyi 2001: 148)

Another enclitic has the form =*oo* and expresses a certain fear that something has happened (Hu Zengyi 2001: 157).

(316) Oroqen (Chaoyang)

*tari jabʊ-tʃaa=oo?*

3SG go-PST=Q

'Is (s)he going?' (Hu Zengyi 2001: 157)

Quite clearly, this is a loan from Mongolian =*UU* that may have acquired a special semantics in Oroqen. Alternative questions may be marked with the enclitic =*jɔɔmaa* ~ =*jooməə* that is of Mongolic origin and may combine with a cognate of Evenki =*Ku*.

5 Survey of the grammars of questions in Northeast Asia

(317) Oroqen (Chaoyang)

a. *əri bajta tǎdzǎǎ joomǎǎ, olook joomǎǎ?*  
 this matter real Q false Q

‘Is this matter true or false?’

b. *əri bajǎ vnaadzı-tfi jǎmmaa.gvǎ, utǎ-tfi joomǎǎ.guu?*  
 this person girl-POSS Q boy-POSS Q

‘Does he (this man) have a girl or a boy?’ (Hu Zengyi 2001: 158)

Apparently, Oroqen also has borrowed the Chinese marker *ba* 吧, but sometimes has two vowel harmonic variants *baa* and *bǎǎ*. It has a long vowel as in some varieties of Khorchin, Dagur (§5.8.2), and Solon (see below).

(318) Oroqen (Chaoyang)

*əri mǎrin aja mǎrin baa?*  
 this horse good horse Q

‘This horse is a good one, right?’ (Hu Zengyi 2001: 157)

Oroqen has also borrowed the Chinese interrogative disjunction *háishi* 还是 ‘or.Q’ for alternative questions. As in Chinese, no additional question marker is present in this example.

(319) Oroqen

*yabu-fa hafı yǎ-fa?*  
 walk-PST or what-PST

‘Did you go or what?’ (Li Fengxiang 2005: 56)

A slightly different picture can be drawn for the Xunke dialect of Oroqen, which has a large amount of question markers (Zhang Yanchang, Li Bing, et al. 1989: passim). The enclitic *=jE* marks polar and, optionally, content questions, which makes a connection to Dagur clear. One of their examples given is an alternative question that contains the two markers *=jǎ* and *=jǎ*. These must be vowel harmonic variants of *=jE*. Thus, the enclitic is even more similar to some subdialects of Dagur that also exhibit vowel harmony in this form. Xunke Oroqen has likewise borrowed the markers *=ɔ* (expressing doubt) from Mongolian *=UU*, and perhaps *baa* ~ *bǎǎ* from Chinese *ba* 吧. Alternative questions may either be marked twice with one of the two markers *ɔmal* and *jǎma* or may take a disjunction *aaki* that may either stand alone or may be combined with other question markers. The origin of *ɔmal* is unclear but possibly may be treated as a variant of *jǎma*. Furthermore, there is a tag question marker *unti*, which looks somewhat similar to the negative copula in Solon and Oroqen that developed out of an adjective meaning ‘different’ (Hölzl 2015a). However, in Xunke Oroqen, the forms are *oŋto* ‘NEG’ and *wǎntǎ* ‘different’ (Zhang Yanchang, Li Bing, et al. 1989: 183).

(320) Oroqen (Xunke)

a. *nǎnıń tǎrǎki=jǎ waa-tǎa aaki gujtǎn=jǎ?*  
 3SG boar=Q kill-PST or roe.deer=Q

‘Did (s)he kill a boar or a roe deer?’

- b. *cii i.rə ɣənə-ni=jɛ?*  
 2SG where go-2SG=Q  
 ‘Where are you going?’
- c. *bii umun.nə ɣənə-tɕə-w, unti?*  
 1SG.NOM once go-PST-1SG right  
 ‘I have been there once, right?’ (Zhang Yanchang, Li Bing, et al. 1989: 123, 126, 131)

In **Huihe Solon**, there is an enclitic =*gi(i)*, which is accompanied by an additional rising intonation. It marks polar and alternative questions.

(321) Solon (Huihe)

- a. *eri üxür aya=gii?*  
 this ox good=Q  
 ‘Is this ox good?’ (Tsumagari 2009a: 7)
- b. *fɪ.n-i bəj-fi aja=gi, əru=gi?*  
 2SG.OBL-GEN body-2SG.POSS good=Q bad=Q  
 ‘Are you well (or sick)?’ (Chaoke D. O. 2009: 316)

Despite functional, formal, and distributional similarities, Solon =*gi(i)* and Evenki =*Ku* are probably not direct cognates of each other because there is no sound law that would justify the different vowel qualities (e.g., Benzing 1956; Doerfer 1978b). Maybe it is a loan from a Mongolic language, e.g. Buryat =*gü* (§5.8.2). Content questions usually do not show any marking.

(322) Solon (Huihe)

- sii ii.lee tegeji-ndi?*  
 2SG where live-PRS.2SG  
 ‘Where do you live?’ (Tsumagari 2009a: 15)

Like Oroqen, Solon also has a marker *baa* with a long vowel that must derive from Chinese *ba* 吧 and a form *yeeme* that, similar to Khorchin Mongolian *jimee*, can also appear in content questions (§5.8.2).

(323) Solon (Huihe)

- a. *ta.ri üli-see baa?*  
 3SG go-PST Q  
 ‘He went, didn’t he?’
- b. *eri si oxon yeeme?*  
 this COP what Q  
 ‘What is this?’ (Tsumagari 2009a: 15)

5 Survey of the grammars of questions in Northeast Asia

As in Oroqen, alternative questions appear to preserve a cognate of Evenki =*Ku*. Consider the following negative alternative question.

- (324) Solon (?Huihe)  
*fii mörn-fi bajə=guu, aafin=gʋʋ?*  
 2SG horse-2SG.POSS man=Q NEG=Q  
 ‘Do you have horses or not?’ (Hu Zengyi & Chaoke D. O. 1986)

There is limited information on other dialects of Solon, especially the **Ongkor** dialect formerly spoken in Xinjiang. However, there apparently were morphosyntactically unmarked questions that probably had a special intonational contour, e.g. *ši mandii?* ‘Are you strong?’ (Aalto 1979: 11) In addition, there are two forms =*ii* and =*uu*, both of which are probably loans from Mongolian =(y)*ii* ~ =(y)*UU* (§5.8.2). In Even, Negidal, and Uilta there are markers similar to =*ii* (see below). Content questions remain unmarked.

- (325) a. Solon (Ongkor)  
*ər uktu ulu-r uktu=ii?*  
 this road walk-PTCP road=Q  
 ‘Is this road the road (usually) travelled?’  
 b. *baxuu-dže=uu e-dže=uu?*  
 find-PST=Q NEG-PST=Q  
 ‘Was it found or not?’  
 c. **jam** *iši-ndii?*  
 which.one see-PRS.2SG  
 ‘What do you see?’ (Aalto 1979: 8, 9, modified transcription)

The interrogative *jam* in (325c) is probably a loan from Jurchenic that can also be seen in Nonni Solon as *jemu* (326b, see §5.10.3). There is even less information on the *Nonni* dialect of Solon. Nevertheless, at least some examples have been collected by Ivanovskij (1982 [1894]). One dubious example of a negative alternative question apparently relies on juxtaposition. Several content questions remained unmarked as well, and an optional polar question marker has the form =*gi*.

- (326) Solon (Nonni)  
 a. *ši.n-i enin šamine žu-de bi-si-ŋ=gi?*  
 2SG.OBL-GEN mother ?father house-LOC COP-PRS-3=Q  
 ‘Are your parents at home (still alive)?’  
 b. **jemu** *gerbi-či?*  
 which name-2SG.POSS  
 ‘What is your name?’ (Ivanovskij 1982 [1894]: 1)

Except for Oroqen, **Negidal** is probably the most aberrant Ewenic language with respect to question marking. At first glance, the situation is similar to Evenki as there is

a marker that is cognate with =*Ku*. Note the absence of the consonant in the form =*ʊʊ*, which, like Ongkor Solon =*uu*, is quite similar to Mongolian. There are also unmarked polar questions that probably have an intonation similar to Evenki.

(327) Negidal

- a. *noŋan naa.bəjə-ni=ŋuu, naa-ni=ʊʊ?*  
 3SG PN-3SG.POSS=Q PN-3SG.POSS=Q  
 ‘Is he a Negidal or a Nanai?’
- b. *sii saa-s?*  
 2SG know-2SG  
 ‘Do you know?’ (Kazama 2002a: 80, 65)

But Khasanova & Pevnov (2003: 10) mention a morphological marking of questions in Negidal as in the following example. Incidentally, the example also contains a further question marker =*i* (cf. example 327a from Even above).

(328) Negidal

- ii-ǰə-m=i?*  
 enter-FUT.Q-1SG.Q=Q  
 ‘Shall I come in?’ (Kazama 2002a: 127)

According to them, the interrogative future differs from the general future in two points. First, the interrogative future has a short vowel as opposed to the plain future. Second, a different personal ending is employed (e.g., 1SG *-m* instead of *-v*). Compare the following pair of sentences:

(329) Negidal

- a. *eeva iche-ǰa-m?*  
 what see-FUT.Q-1SG.Q  
 ‘What will I see?’
- b. *oǰa-va iche-ǰee-v.*  
 track-ACC see-FUT-1SG  
 ‘I will see the tracks.’ (Khasanova & Pevnov 2003: 10)

The morphological interrogative marking is found in polar, content, as well as in alternative questions and can combine with interrogative enclitics. Consider the following open alternative question with both morphological and enclitic markers.

(330) Negidal

- mozhet bolotki bi-ǰə-m=ŋu ee-ǰa-m=ŋu?*  
 maybe autumn be-FUT.Q-1SG=Q what-FUT.Q-1SG=Q  
 ‘Is it perhaps already autumn or what?’ (Kazama 2002a: 114)

Previous descriptions of Negidal apparently did not mention this interesting feature (see Kazama 2002a: 107, 114, 115). According to Ikegami (1985), in Tungusic languages there are generally two different sets of personal endings (Table 5.111). In Negidal, Set 1 is used after past forms in *-čaa* as well as future forms in *-ja(-ηaa)* and also has a possessive function with nouns. Set 2, on the other hand, can be found after present stems in *-ja* or underived stems. Ikegami (1985: 91) also notes that, according to Kolesnikova & Konstantinova, the future ending *-ja* takes the first person inclusive marker *-p* instead of *-t*. This might indicate a confusion resulting from the interrogative marking and may show that Khasanova & Pevnov's (2003) assumptions are correct.

Table 5.111: Personal endings in Negidal according to Ikegami (1985: 88f.), from Cincius, adjusted

	Set 1	Set 2
1SG	<i>-v</i>	<i>-m</i>
2SG	<i>-s</i>	<i>-s</i>
3SG	<i>-n</i>	<i>-n</i>
1PL.INCL	<i>-t</i>	<i>-p</i>
1PL.EXCL	<i>-vun</i>	<i>-vun</i>
2PL	<i>-sun</i>	<i>-sun</i>
3PL	<i>-tin</i>	-

Accordingly, Set 2 would additionally be used in interrogatives, while Set 1 is found in declarative sentences. There does not appear to be any further description of this phenomenon for Negidal, or for any other Tungusic language for that matter. However, a possible areal connection can be found in Yukaghiric (§5.14.2). As in Negidal, the Yukaghiric interrogative suffixes are restricted to the first person (singular *-m*, plural *-uok ~ -ook*). But the connection to Yukaghiric is not without its problems. First of all, Yukaghiric languages are spoken several thousand kilometers north of Negidal and in Yukaghiric the special interrogative suffixes are only found in content questions. Furthermore, Yukaghiric lacks any special interrogative tense markers. But as specified in §2.14 we may assume that Yukaghiric was once spoken in a much larger territory and that its speaker probably migrated northward along the Lena river from an earlier location close to Lake Baikal, which reduces the distance to the Negidal. But even if the areal connection turns out to be wrong, we are dealing with an interesting typological parallel in which interrogative agreement marking is mostly restricted to the first person and the third person plural remains unmarked.

In **Udihe** polar questions can be marked by intonation only, which is said to be higher and somewhat longer than that of declarative sentences (Nikolaeva & Tolskaya 2001: 807). An element may be moved to a focus position, typically in front of the verb, which is different from Evenki as seen above (Nikolaeva & Tolskaya 2001: 841).

- (331) Udihe  
*uti nii ḡənəə-ni bi bagdəə-mi bua-la?*  
 that man go.PST-3SG 1SG.NOM born.PST-1SG place-LOC  
 ‘Has *he* (the man) gone to my birthplace.’ (Girfanova 2002: 41)

An alternative is the use of an enclitic =*nu* ~ =*gu*, cognate of Evenki =*Ku*, that attaches to the verb in polar questions and to the the element in focus in focus questions. As in Evenki, the scope of the marker also encompasses alternative questions.

- (332) Udihe
- a. *Iwana zugdi-du bi-s’e=nu?*  
 Ivan house-LOC COP-PFV=Q  
 ‘Is Ivan at home?’
  - b. *si bagää-za=nu bi-s’e-i?*  
 2SG other.side-N-Q COP-PFV-2SG  
 ‘Were you on the other side of the river?’
  - c. *xeleba bie=nu anči=nu?*  
 bread be.PRS.HAB=Q NEG=Q  
 ‘Is there bread or not?’ (Nikolaeva & Tolskaya 2001: 809, 812)

Content questions do not normally take any morphosyntactic marking.

- (333) Udihe  
*j’e-le ṅansule-i?*  
 which-LOC study-2SG  
 ‘Where do you study?’ (Nikolaeva & Tolskaya 2001: 801)

The semantic scope of =*nu* is thus identical to Evenki, but Udihe has a further enclitic =*nA* that has a contrastive function. It remains dubious whether this form has any connection with Manchu =*nA* (see below).

- (334) Udihe  
*xuda=na?*  
 fur=Q  
 ‘And what about the *fur*?’ (Nikolaeva & Tolskaya 2001: 808)

The enclitic is also often used together with an interrogative word. Within the following question the contrastive focus lies not on the interrogative but on the river.

- (335) Udihe  
*ei=ne j’e.u bāsa-ni?*  
 this=Q what river-3SG  
 ‘And what is *this* river called?’ (Nikolaeva & Tolskaya 2001: 808)

Alternative questions may also be formed with *-(e)s(i)* of unknown origin. In example (336a) a content question is followed by an alternative question (§4.4).

(336) Udihe

- a. *j'e-we xokto-ni, kãŋaa-s ogböö-s?*  
 what-ACC footstep-3SG.POSS deer-Q elk-Q  
 'Whose footsteps are these, a deer's or an elk's?'
- b. *iwana sin-du kusige-we bu-ge, e-si-ni bu-oo-s(i)?*  
 PN 2SG.OBL-DAT knife-ACC give-PFV neg-PST-3SG give-PST-Q  
 'Has Ivan given you the knife or not?' (Nikolaeva & Tolskaya 2001: 811, 255)

This latter construction is probably not a tag question construction but an alternative question with a question marker on the second alternative only, which is also attested for Kilen and Manchu.

According to Nikolaeva & Tolskaya (2001: 351), there are tag questions that are formed with the help of the interrogative *j'e.u* 'what', which may be attributed to Russian influence (cf. §5.5.2.2).

(337) Udihe

- em'e-i, j'e.u?*  
 come.PFV-2SG what  
 'You came, didn't you?' (Nikolaeva & Tolskaya 2001: 351)

Udihe and Oroch have a very interesting open alternative question construction in which the second alternative is an inflected interrogative verb. This pattern has been adopted by Kilen from Udihe. We have already observed a similar construction in Oroqen (319), but with the Chinese disjunction instead of double marking.

(338) Udihe

- su xulisee-u=nu jaa-u=nu?*  
 2PL go.PST-2PL=Q what.-2PL=Q  
 'Did you go, or what?' (Nikolaeva & Tolskaya 2001: 811)

(339) Kilen

- su ənə-xəi=nə ja-xəi=nə?*  
 2PL go-PERF=Q what-PERF=Q  
 'Did you go or what?' (Zhang 2013: 158)

(340) Oroch

- agduči-za-i=nu jaa-za-i=nu?*  
 tell-FUT-1SG=Q what-FUT-1SG=Q  
 'Should I tell or what?' (Tolskaya & Tolskaya 2008: 98, from Avrorin)

Example (340) from **Oroch** shows the interrogative enclitic which, similar to Nanai and Udihe, has the form =*nu* and is optional in polar questions. Content questions remain unmarked.

## (341) Oroch

- a. *sii čihala-i-si?*  
2SG agree-PRS-2SG  
'Do you agree?'
- b. *sii jaa-va xuana-i-si?*  
2SG what-acc argue-PRS-2SG  
'What do you argue about?'
- c. *sii amba bi-si=nu?*  
2SG evil.spirit COP-PRS=Q  
'Are you an evil spirit (the devil)?' (Avrorin & Boldyrev 2001: 184)

In 1958, a team of unknown scientists from China gave a handful of comparative word lists for five Tungusic languages in China. Their list also contains two sentences that can be translated as 'when do you come back?' (多怎回来) and 'where do you go?' (NDSSLD 1958: 82). Unfortunately, they transcribed all languages with the help of Chinese characters, which makes the analysis less easy. Additionally, some characters were written incorrectly. The following gives the corrected sentences in Chinese transcription and its rendering in official Pinyin spelling followed by a rough approximation of the original languages. The transcription, analysis and glossing is mine. Interestingly enough, the set of languages is not completely identical to the five officially recognized languages today. There are no sentences from Sibe, but from **Hezhen** (*hèzhēn* 赫真) which refers to the dialect of Nanai spoken in China. Hezhen is not very well known (cf. An Jun 1986: 79–86) and probably extinct by now, while *Kilen* (*qíléng* 奇楞) has been described in several grammatical sketches. The Hezhen data are thus potentially very important. Both Hezhen and Kilen are classified together as the Hezhe (*hèzhé* 赫哲) language and are treated as dialects by the authors of NDSSLD (1958). Of the five languages only Hezhen and Kilen are included here for the sake of brevity.

## (342) Hezhen

- a. 哈利吉朱衣西 [ha li ji zhu yi xi]  
*xali dži-dzu-i-ci?*  
when come-REGR-PRS-2SG  
'When do you come back?'
- b. 好西額奴衣西 [hao xi e nu yi xi]  
*xaoci enə-i-ci?*  
where.to go-PRS-2SG  
'Where do you go?' (NDSSLD 1958: 82)

(343) Kilen

- a. 阿黑額莫土衣西 [a hei<sup>25</sup> e mo tu yi xi]  
**ali** emə-tu-i-ci?  
 when come-REGR-PRS-2SG  
 ‘When do you come back?’
- b. 鴨勒額訥衣西 [ya le e ne yi xi]  
**ya-le** ene-i-ci?  
 which-ALL go-PRS-2SG  
 ‘Where do you go?’ (NDSSLD 1958: 82)

In both languages content questions do not take any morphosyntactic marking. As will be further explained in §5.10.3, Kilen interrogatives exhibit affinities with Udegheic, which explains the absence of the initial consonant in *ali*, as opposed to Hezhen *hali* ‘when’, and the interrogative *yale*, instead of Hezhen *xaosi* ‘whither’ (Udihe *ali*, *j’ele*, Nanai *xaali*, *xaosi*).<sup>26</sup> Schmidt (1928b: 241) mentions the Samar sentence *xajadži džidžisi?* ‘Where did you come from?’ Samar is not very well-known, but is clearly very similar to Nanai as well (e.g., Nanai *xajaji* ‘whence’).

There are several descriptions of Kilen that differ more or less strongly from each other. According to Zhang (2013: 157f.), **Kilen** expresses polar questions with rising intonation on the last word of the sentence, e.g. *ci sa?* ‘Do you know?’. However, Kilen was heavily influenced by Chinese, in fact, Chinese may by now have replaced Kilen completely, leaving Kilen extinct. Following to Zhang (2013: 158), Kilen borrowed the three interrogative particles *ba* 吧, *ma* 吗, and *(y)a* 啊/呀, all of which are possible in the sentence above, e.g. *ci sa=a?* ‘Do you really know?’. Most likely, =*a* is not of Chinese origin, however. Several examples of polar questions in Zhang Yanchang, Zhang Xi, et al. (1989) were either unmarked (showing rising intonation) or marked with the final question marker =*a*. Note that it never followed anything but the second person singular agreement form -*ci* and was always written attached to it. Nevertheless, it is better analyzed as enclitic =*a* that may appear in both polar and content questions, which might speak instead in favor of a connection with Manchu =*o*.

(344) Kilen

- a. *ci arki-wə ɔmi-mi bi-ci=a?*  
 2SG alcohol-ACC drink-CVB COP-2SG=Q  
 ‘Are you drinking alcohol?’
- b. *ia-lə ənə-ji-ci=a?*  
 which-ALL go-PRS-2SG=Q  
 ‘Where are you going?’ (Zhang Yanchang, Zhang Xi, et al. 1989: 87f.)

<sup>25</sup>The character *hei* 黑 should instead read *li* 里.

<sup>26</sup>Both Hezhen and Kilen show characteristics that suggest a basic connection to Nanai, e.g. the absence of an initial consonant in *ene* ‘to go’ (Nanai *ənə*-, Udihe *ɲene*-, Manchu *gene*-). The verb *emə* ‘to come’ in Kilen was most likely borrowed from Udihe (Nanai *ji*-, Udihe *eme*-, Manchu *ji*-).

Note the absence of the marker in the otherwise identical sentence (47b) above. An Jun (1986) already mentions examples with the Chinese enclitic *ba* 吧. In his data, *a* is not written attached to the preceding word and can also follow forms other than the second person singular.

- (345) Kilen  
*ci.n-i            agə-ci            biχan            fuli-m            ən-χ-ni=a?*  
 2SG.OBL-GEN e.brother-2SG.POSS wilderness hunt-CVB go-PST-3SG=Q  
 ‘Did your elder brother go to hunt?’ (An Jun 1986: 36)

In alternative questions the Chinese interrogative disjunctive *háishi* 还是 ‘or.Q’ may be employed but is combined with the marker =*a*.

- (346) Kilen  
*ci   əi-wə    xəci ti-wə    gələ-ji-ci=a?*  
 2SG this-ACC or.Q that-ACC ?want-PRS-2SG=Q  
 ‘Do you want this or that?’ (Zhang Yanchang, Zhang Xi, et al. 1989: 45, simplified)

Unlike other alternative question constructions among Tungusic languages, the question marker appears only once and does not attach to the elements in focus.

A further enclitic called a “contrastive particle” by Zhang (2013: 159) seems to have been borrowed from Udihe and marks polar and focus questions.

- (347) Kilen  
 a. *ci   adəqəli=nə?*  
    2SG cold=Q  
    ‘Are you cold (or not)?’  
 b. *sursarə=nə talaxa?*  
    tasty=Q    grilled.fish  
    ‘Is the grilled fish tasty (or not)?’ (Zhang 2013: 159)

As seen in example (365) above, it also marks alternative questions. Most likely it has been borrowed from Udihe =*nu*, but it exhibits certain similarities to Udihe =*nA* as well.

Nanai is the best described language from the Nanaic branch and there is even a good description of question intonation by Baitchura (1979: 294) (underlining removed).

In general questions, the tone movement in the vowel of the final syllable has a clearly and strongly manifested rising character, whereas the mean and the maximal tone heights surpass those of the preceding vowels in cases in which no interrogative particle is present in the sentence. If there is such a particle (e.g., *nu*), the rise of the tone at the end of the sentence is not so high, its pitch being a little lower in comparison to the tone heights of vowels at the beginning of the sentence.

5 Survey of the grammars of questions in Northeast Asia

There are sentences with and without the enclitic. In 1858, Venjukov recorded a polar question without enclitic among the Ussuri Nanai.

- (348) Nanai (Ussuri)  
*anda, duman bira goró?*  
 friend PN river far  
 ‘Friend, is the Duman river far away?’ (Alonso de la Fuente 2011: 14, from Venjukov)

Similar to Evenki or Udihe, the enclitic =*nu* in Nanai does not appear in content questions (which remain unmarked), but marks more than one question type, including polar, alternative, and possibly focus questions.

- (349) Nanai (Najkhin)
- a. *swə aja-so=nu?*  
 2PL good-2PL.POSS=Q  
 ‘Are you all well?’
  - b. *swə sogdata-wa wanda-mari maŋbo-ci=nu namo-ci=nu ənə-i-su?*  
 2PL fish-ACC take-CVB.PL river-ALL=Q sea-ALL=Q go-PRS-2PL  
 ‘Do you go to the river or sea to catch fish?’
  - c. *si xali əusi ji-ci-si?*  
 2SG when here come-PST-2SG  
 ‘When did you come here?’ (Ko & Yurn 2011: 155, 68, 52)

Regarding the last sentence compare example (370) from Hezhen above.

In **Ulcha** the enclitic marks focus, alternative, and (optionally) content questions. Such an extension of scope can also be observed in Mongolian (§5.8.2).

- (350) Ulcha
- a. *lʷča gasa=nʷʋ?*  
 PN village=Q  
 ‘(Is it) a Russian village?’
  - b. *xasu aňan=nʷʋ, nadan aňan=nʷʋ ŋuŋgun aňan=nʷʋ bi-či-ni?*  
 how.many year=Q seven year=Q six year=Q COP-PST-?3SG  
 ‘How many years has it been, seven years (or) six years?’ (Kazama 2002b: 79, 86)
  - c. *saŋxai, xai.mi soŋg-i?*  
 PN why cry-?PRS  
 ‘Sanghai, why are you crying?’ (Schmidt 1923b: 235f.)

Within Nanaic, **Uilta** has the most interesting marking of questions. Polar questions in Uilta have both rising intonation and an interrogative clitic =(y)i that might be related to the one found in Even, Negidal, and Ongkor Solon, although these were perhaps

borrowed from Mongolic. It always follows the verb (Ptryk Czerwinski p.c. 2018). In addition, there is a specialized marker =*ga* ~ =*ka* for content questions that cannot be found in any other Tungusic language.

- (351) Uilta (Southern)
- a. *eri aya ulaa=yi?*  
 this good reindeer=Q  
 ‘Is this a good reindeer?’
- b. *ɲui sinda-xa-ni=ga?*  
 who come-PST-3SG=Q  
 ‘Who came?’ (Tsumagari 2009b: 15)

Apparently, the marker in content questions is not obligatory as there are also several examples without it.

- (352) Uilta (Northern)
- khoni bi-si sii?*  
 how COP-PRS 2SG  
 ‘How are you?’<sup>27</sup> (Funk 2000: 150)

The Uilta polar and content question markers can almost certainly be attributed to influence from Amuric (see Sections 3.1 and 5.2.2). Within Ikegami’s (1997) dictionary there are not only examples with =*ga*, but also with a marker =*gəə*.

- (353) Uilta
- tari nari ɲui=gəə?*  
 that person who=Q  
 ‘Who is that person?’ (Ikegami 1997: 145)

However, riddles recorded by Ikegami contain yet another variant with a final -*k*.

- (354) Uilta (Southern)
- xai=gəək?*  
 what=Q  
 ‘What is (this)?’ (Ikegami 1958: 93)

The origin of the final -*k*, which can also appear in children’s games, remains partly unclear and other examples with similar constructions exhibit the marker =*ga*, instead.

- (355) Uilta (Southern)
- eri xai=ga?*  
 this what=Q  
 ‘What is this?’ (Tsumagari 2009b: 15)

<sup>27</sup>Regarding the use of the interrogative, cf. Russian *kak dela?*/как дела?

Patryk Czerwinski (p.c. 2018) was so kind to check with some of the last speakers of the northern dialect. According to his fieldwork, the variant with =gəək is still used as an “embellishment” of questions and is marked with respect to the other variants. *xai tari?*, *xai=ga tari?*, and *xai=gəək tari?* are said to have more or less the same meaning ‘What is that?’.

Nakanome (1928: 21, 50) already mentioned two different forms, the unproblematic form <ga> and yet another variant written as <nö> that was probably pronounced with a velar nasal [ŋ] and a vowel quality comparable to the form =gəə recorded by Ikegami, i.e. [ŋə] (see §3.1).

- (356) Uilta  
*hai-wö gade-si=ñö?*  
 what-ACC buy-2SG=Q  
 ‘What do you (want to) buy?’ (Nakanome 1928: 52)

In addition, there are variants with a fricative in intervocalic position, e.g. [ŋui=ɣə], [ŋui=ɣə(ə)k] ‘who-Q’ (Patryk Czerwinski p.c. 2018). Most likely, we are dealing with one enclitic that undergoes both vowel harmonic and consonant alternations depending on the preceding syllable (i.e. =KA(A)). In my eyes, Nivkh =ŋa is the most likely source of this enclitic in Uilta (see §3.1).

It is an open question whether =gəək is an independent form or a variant of =KA(A). A -k can also appear in answers to riddles and might be a suffix. However, the form =gəək apparently exhibits no vowel harmony and only appears in special contexts, which might suggest that it is in fact a different form (Patryk Czerwinski p.c. 2018).

In the northern dialect, the question marker seem to be more strongly fused with the preceding elements (-čee < -či ‘3PL’ + =KA ‘Q’).

- (357) Uilta (Northern)  
*xooni to-li-čee?*  
 how do-P.FUT-3PL.Q  
 ‘How are they doing?’<sup>28</sup> (Yamada 2016: 192)

No examples for alternative, focus, and tag questions have been found in the relevant literature (e.g., Ikegami 2002). According to Patryk Czerwinski (p.c. 2018), focus questions do not show any difference with respect to polar questions. He elicited the following two alternative questions for me. The analysis roughly follows Tsumagari (2009b).

- (358) Uilta (Northern)  
 a. *sii xoo-tai ŋəŋnee-si, oskoola-tai yyuu, duku-takki yyuu?*  
 2SG which-DIR go.PRS-2SG school-DIR Q house-DIR.REFL.POSS Q  
 ‘Where are you going, are you going to school or to your house?’

<sup>28</sup> どうすればよいでしょう? in Japanese. For the use of the interrogative, cf. Russian *kak dela?*/как дела?

- b. *sii oskoola-tai ɲəpɲee-si=yi, duku-takki yyyuu?*  
 2SG school-DIR go.PRS-2SG=Q house-DIR.REFL.POSS Q  
 ‘Are you going to school? Or to your house?’
- c. *xai-wa dəptu-li-ssee, sundattaa, ulissə yyyuu?*  
 what-ACC eat-FUT-2SG.Q fish.ACC meat.ACC Q  
 ‘What will you eat, fish or meat?’

In the second example, there is the polar question marker =(y)i at the verb in the first alternative. The second alternative takes what appears to be a question marker *yyuu*. In the first example, because of the ellipsis of the verb, the marker *yyuu* is found on each alternative. In the third example, the marker *yyuu* only appears on the second and last alternative. The preceding content questions exhibits a fused question marker similar to the one seen before (perhaps *-si + =KA > -see*). The only possible question tag in Uilta is *ii* ‘yes’ (similar to Russian), although this is difficult to identify, given the formal resemblance with the polar question marker =*yi(i)* (Patrik Czerwinski p.c. 2018).

Question marking similarly aberrant to that in Uilta can be observed in the entire Jurchenic branch, but especially in Written Manchu. For **Manchu**, book three of the *Qingwen Qimeng* (Wuge Shouping & Cheng Mingyuan 1730, translated by Wylie 1855) lists a number of interrogative forms: *na, ne, no, nu, ya*, all of which are probably enclitics. The first three must be vowel-harmonic variants of one form =*nA*, which is similar to Udihe, although a connection remains doubtful. The enclitic =*nu* may be cognate of Evenki =*Ku*, Udihe =*nu*, and Nanai =*nu*, but is not often encountered.

(359) Manchu

- a. *waka-ra-rakū=na?*  
 mistake-V-P.IPFV.NEG=Q  
 ‘Will (you) not blame (me) then?’
- b. *gene-rakū=ne?*  
 go-P.IPFV.NEG=Q  
 ‘Will (you) not go?’
- c. *o.jo-rakū=no?*  
 become-P.IPFV.NEG=Q  
 ‘Will it not do?’
- d. *gisu-re-rakū=nu?*  
 word-V-P.IPFV.NEG=Q  
 ‘Will (you) not speak?’ (Wuge Shouping & Cheng Mingyuan 1730; Wylie 1855: 171)

The marker =*ya* is also not very frequent and, might have been borrowed from Chinese =(y)a 啊/呀, e.g. *inu=ya?* ‘Is it so?’ It appears that =*ni* and =*o* are not only the most neutral but also the most frequent question markers. There is not much information about the two enclitics, but both appear in polar, alternative, and content questions, which makes

5 Survey of the grammars of questions in Northeast Asia

Manchu different from most other Tungusic languages, but not Oroqen. The enclitic =*ni* has the reduced form *n* after the existential negator *akū*.

(360) Manchu

- a. *ere gese baita geli bi=ni?*  
 this like matter also COP=Q  
 ‘Was there ever anything like this?’
- b. *ere ai turgun=ni?*  
 this what reason=Q  
 ‘What is the reason for this?’
- c. *si ere bithe=be hola-ki se-mbi=o?*  
 2SG this book=ACC read-OPT say-IPFV=Q  
 ‘Do you want to read this book?’
- d. *geren niyalma-i dorgi fala-ci aca-ra-ngge we-be=o?*  
 all person-GEN inside punish-COND meet-IPFV-N who-ACC=Q  
 ‘Who among all men is to be punished?’ (Wuge Shouping & Cheng Mingyuan 1730; Wylie 1855: 131, 133, 140)
- e. *si min-de bu-mbi=o bu-rakū.n?*  
 2SG 1SG.OBL-DAT give-IPFV=Q give-IPFV.NEG.Q  
 ‘Are you going to give it to me, or not?’ (Hauer 2007: 67, from the *Jinpingmei*)

Possibly influenced by *akūn*, the words *sain* ‘good’, *tašan* ‘false’, and *yargiyan* ‘true’ have the special interrogative forms *saiyūn*, *tašun*, and *yargiyūn*. The last example (360e) consists of a negative alternative question in which, unlike any other Tungusic language except Uilta, two different question markers may be employed. In Manchu, there is a wealth of such verb doubling constructions for questions in which the second verb is always negated (Table 5.112). Only one of these patterns marks both verbs with a question marker and two do not have any marker at all, which may be due to Chinese influence. In one case, the second alternative takes two markers. In most cases, there is one marker found at the second negated verb (cf. Kilen and Udihe above).

Given its semantic scope and the possibility that =*o* alone may mark a negative alternative question, a connection to Kilen =*a* seems possible. There are further constructions not mentioned by Gorelova, which have no reduplication of the verb.

(361) Manchu

- a. *gebu ali.bu-ha=o akūn?*  
 name submit-P.PVF=Q NEG.Q  
 ‘Have (you) enrolled (for the exam) or not?’ (von Möllendorff 1892: 28)
- b. *ere kemuni tolhin waka=o se-me=o?*  
 this still dream NEG=Q say-CVB.IPFV=Q  
 ‘This isn’t a dream or is it?’ (Di Cosmo 2006: 87, 104, 131)

Table 5.112: Negative alternative question patterns in Manchu (Gorelova 2002: 325f.); *-mbi* 'IPFV', *-rA* 'P.IPFV', *-hA* 'P.PFV', *-hAkū/-rakū* 'NEG', *se-* 'to say', *bi-* 'COP'

V1	V2	AUX
V-ra	V-rakū	
V-ra	V-rakū= <b>n</b>	
V-ra	V-rakū	bi= <b>o</b>
V-ra	V-rakū	se-me= <b>o</b>
V-ra	V-rakū= <b>nio</b>	
V-ha	V-hakū	
V-ha	V-hakū= <b>n</b>	
V-ha	V-hakū= <b>n</b>	bi-he= <b>o</b>
V-mbi	V-rakū	
V-mbi	V-rakū= <b>n</b>	
V-mbi= <b>o</b>	V-rakū= <b>n</b>	
V-mbi-he	V-mbi-hekū	
V-mbi-he	V-mbi-hekū= <b>n</b>	

As opposed to the previous constructions, this last example (361b) has the same question marker used twice, which may be due to the presence of the negative copula *waka*, after which apparently only =*o* can be found. Aixinjueluo Yingsheng (1987a: 72) argued that a Mandarin interrogative construction with sentence-final *yǒu ma?* 有吗 'EX Q', apparently found in the Peking dialect, is a calque of Manchu *bi=o?* 'EX=Q'.

It is often claimed that the two markers =*ni* and =*o* may also be attached behind one another to form the complex marker =*nio*.

(362) Manchu

*ere sain akū=nio?*

this good NEX=Q

'Isn't this good?' (Wuge Shouping & Cheng Mingyuan 1730; Wylie 1855: 134)

This would be a very unusual pattern among Tungusic languages. But apart from this analysis into two question markers, which is a rather unexpected, there is a more plausible explanation that treats =*nio* as one marker that was borrowed from Korean (see below). Also remember that, following *akū*, the marker =*ni* usually takes the form *-n*.

In *Sibe*, polar questions are regularly expressed with the enclitic =*na* that seems to correspond to the Manchu form =*nA* above but does not exhibit vowel harmony. It marks polar and alternative questions. In both polar and content questions, there is sometimes an element =*jə* that might correspond to *Dagur* =*yee*. But its status as a question marker remains rather dubious. Like many languages in China, *Sibe* has adopted the Mandarin question marker *ba* 吧.

(363) Sibe

- a. *ʃi guldza-ji nan=na?*  
 2SG PN-GEN person=Q  
 ‘Are you from Yining?’
- b. *ʃi evʃi=jə?*  
 2SG how=?Q  
 ‘How are you?’
- c. *tʃəhsə avqa-ji arvun evʃi?*  
 yesterday sky-GEN form how  
 ‘How was the weather yesterday?’ (Chaoke D. O. 2006: 207, 280, 277)
- d. *ə.r mi.n dzaq=na, ʃi.n dzaq=na?*  
 this 1SG.GEN thing=Q 2SG.GEN thing=Q  
 ‘Is this mine or yours?’
- e. *so.n tə.va mori<sup>n</sup> gum ambu ba?*  
 2PL.GEN there horse all big Q  
 ‘In your place all horses are big, aren’t they?’ (Zikmundová 2013: 49, 95)

In general, there are very few descriptions of possible tag questions in Tungusic languages. Sibe is somehow exceptional because at least two different tag question patterns were recorded.

(364) Sibe

- a. *tə-s ganbi ʃi.n-i-ŋə vaq, məndzəŋ=ba?*  
 that-PL pencil 2SG.OBL-GEN-N NEG right=Q  
 ‘Those aren’t your pencils, right?’
- b. *tə.r əm nan siv, məndzəŋ=na?*  
 that one man teacher right=Q  
 ‘That one is a teacher, right?’
- c. *ʃi mi.ni-d utʃi-v li-maq ʃinda, o-m=na?*  
 2SG 1SG.OBL-DAT door-ACC open-CVB.PFV stay[IMP] be-IPFV=Q  
 ‘Hold the door open for me, will you?’ (Chaoke D. O. 2006: 91, 102, 343)

The first type could be a calque and partial loan from Mandarin *duì ma/ba* 对马/吧. The latter type with the verb *o-* ‘to become, to be, to be permissible’ (Norman 2013) possibly is a calque of Mandarin *kěyǐ ma* 可以吗 or *xíng ma* 行吗 (§5.9.2.1). There are also parallels in Khorchin Mongolian (§5.8.2).

Records of Sibe from the beginning of the 20th century that were strongly influenced by Written Manchu have been recorded by Muromski. They contain several question markers, =*na(a)*, =*nu(u)* ~ =*nü*, =*ü* ~ =*u*, and =*o* (Kałużyński 1977: 53). The marker =*U* might be of Mongolic origin and is the only one that is unknown from Manchu. It appears to have fused with the imperfective or dictionary form *-mbi* of Written Manchu.

(365) Sibe

*mi.n-i gala-ci tuči-mbü?*

1SG.OBL-GEN hand-ABL come.out-IPFV.Q

‘Can you escape my hands?’ (Kałużyński 1977: 53)

**Sanjiazhi Manchu** also shows some of this variation. The following examples were collected in 1961 and contain the markers =*nə*, =*nu*, and =*ni*. The last one seems to be restricted to content questions that are optionally unmarked, while the other two (=NU) appear in polar questions. Enhebatu treats them as variants of the same form.

(366) Manchu (Sanjiazhi)

a. *ur baiti-bur dɔndzi-γ=nə?*

this matter-ACC hear-P.PFV=Q

‘Did you hear about this matter?’

b. *u.lu ai dza<sup>q</sup>xa=ni?*

this what thing=Q

‘What (kind of thing) is this?’

c. *ai gurvur?*

what name

‘What (is your) name?’ (Enhebatu 1995: 300, 296, 55)

Kim et al. (2008: 45), who did fieldwork in Sanjiazhi in 2005 and 2006, recorded the markers =*no* and =*nə*. They claim that the latter is a loan from Mandarin *呢*. Sanjiazhi has also borrowed Mandarin *吧*. For alternative questions the Chinese disjunction *háishi* 还是 ‘or.Q’ has been adopted.

(367) Manchu (Sanjiazhi)

a. *ci qanqxy-by ba?*

2SG be.thirsty-P.PFV Q

‘You must be thirsty, right?’ (Enhebatu 1995: 55)

b. *najnaj čičikal=dili ji-xə-niηə haishi xaəlbin=dili ji-xə-niηə?*

grandmother PN=ABL come-P.PFV-N or.Q PN=ABL come-P.PFV-N

‘Is your grandmother from Qiqihar or from Harbin?’ (Kim et al. 2008: 214)

Kim et al. (2008: 46) mention an enclitic =*ja* ~ =*jə* that they call an “‘intimacy’ particle”. It may appear in questions but is not restricted to them. A connection to the Sibe and Dagur enclitic seems more likely than with Mandarin *呀*.

The **Yibuqi** dialect of Manchu presents a situation very similar to Sanjiazhi Manchu. The usual question marker has the form =*no*, content questions remain unmarked, and the Mandarin disjunction may be employed in plain and negative alternative questions.

(368) Manchu (Yibuqi)

- a. *pi putA tʂə-mi=no?*  
 1SG.NOM food eat-IPFV=Q  
 ‘(Can) I eat the food?’
- b. *cick‘ə və tci-γə?*  
 yesterday who come-P.PFV  
 ‘Who came yesterday?’
- c. *ci tci-mi xεʂl pi kənə-mi?*  
 2SG come-IPFV or 1SG go-IPFV  
 ‘Do you come or do I go?’<sup>29</sup> (Zhao Jie 1989: 164, 184, 188)

The Yibuqi dialect additionally borrowed the Mandarin polar question marker *ma* 吗.

(369) Manchu (Yibuqi)

- so kəm tci-γə mA?*  
 2PL all come-P.PFV Q  
 ‘Have you all come?’ (Zhao Jie 1989: 154)

Aihui Manchu has the standard polar question marker =no. A form =je similar to Sibe is attested, but its meaning is not perfectly clear. Content questions usually remain unmarked. Alternative questions take the Mandarin disjunction *háishi* 还是 ‘or.Q’.

(370) Manchu (Aihui)

- a. *ci mandzo gisun bəvə.na-m=no?*  
 2SG Manchu language be.able-IPFV=Q  
 ‘Can you speak Manchu?’
- b. *ε-bəri jov-ka?*  
 what-DIR go-P.PFV  
 ‘Where did (she) go?’
- c. *anə gəl məri-m dzi-γə=je?*  
 why again return-CVB.IPFV come-P.PFV=?Q  
 ‘Why have you come back again?’
- d. *ə.rə dzaqa ʂl fə-niŋŋə xεʂl itci-ŋŋə?*  
 this thing COP old-N or.Q new-N  
 ‘Is this thing old or new?’ (Wang Qingfeng 2005: 210, 229, 228, 243)

For Aihui Manchu a tag question different from Sibe has been recorded. It may have been partly calqued from Mandarin *duì bu duì* 对不对 or *duì ba* 对吧.

<sup>29</sup>Here, the disjunction may also have the form *ʂl*.

- (371) Manchu (Aihui)  
*bi agə-dəɾə gəm adzigən, ino vaqa=ba?*  
 1SG.NOM e.brother-ABL all little correct wrong=Q  
 ‘I’m smaller than all (my elder) brothers, isn’t that right?’ (Wang Qingfeng 2005: 236)

It may be noted that both *inu* and *waka* also function as positive and negative one word answers, respectively, in Written Manchu.

The two languages **Bala** and **Alchuka** add important pieces to the puzzle. Both preserves a cognates of Manchu =*o*. Compare the following two sentences.

- (372) ?Bala  
*ci.n nianli ai-və-t’ bi=ɔ?*  
 2SG.GEN washing.hammer what-place-LOC COP=Q
- (373) Lalin/Jing Manchu  
*ci.n-i nijandʒ’a.k’u ai-ba-de bi-x=ɔ?*  
 2SG.OBL-GEN washing.hammer what-place-LOC COP-PFV=Q  
 ‘Where is your washing hammer?’<sup>30</sup> (Mu Yejun 1987: 25)

Alchuka, in addition to =*ɔ*, has a variant =*kɔ* with an unaspirated [k]. This form is related to Manchu =*o* as well, as can be observed from a comparison of Alchuka *ələ-mei=kɔ* ‘fear-IPFV=Q’ (Mu Yejun 1986: 16) with Manchu *gele-mbi=o* (Aixinjueluo Yingsheng 1987b: 15) that were attested in the same sentence. Bala also has a form =*ŋɔ* that is most likely cognate with Sanjiazi =*nɔ*, Aihui Manchu =*no*, Yibuqi Manchu =*no*, and Manchu =*nio*.

- (374) Bala
- a. *ci ənə=ŋɔ?*  
 2SG go=Q  
 ‘Are you going?’
- b. *ci.n amin=ŋɔ?*  
 2SG.GEN father=Q  
 ‘Is it your father?’ (Mu Yejun 1987: 31)

Table 5.113 summarizes interrogative markers in Tungusic languages. Kyakala, Jurchen A, Jurchen B, Kili, and Arman have been excluded for lack of information. To the best of my knowledge, the origin of the Jurchenic question markers have never been described satisfactorily. But given their presence in Jurchenic, exclusively, and the lack of a good internal etymology, a borrowing from a neighboring language seems plausible. I argue that most of them (Manchu =*o*, =*n(i)*, =*nio*, =*nA*) were perhaps borrowed from Koreanic, which had longstanding contacts with Jurchenic. The details are presented in

<sup>30</sup>洗衣棒锤 in Chinese. Norman (2013) translates Written Manchu *nianca-kū* as ‘a wooden stick for beating starched clothes while washing’.

§5.7.2. Manchu =*nu* might be inherited from Proto-Tungusic. Aihui Manchu *-je*, Sanjiazhi Manchu *-jA* as well as Sibe *-jə* may have been borrowed from Dagur. The disjunction in Kilen, Oroqen, and Manchu dialects was borrowed from Mandarin.

Regarding the syntactic behavior of interrogatives in content questions, Malchukov & Nedjalkov (2010: 343f.) offer the following summary.

Question formation need not involve WH-movement in Tungusic languages. For some languages, WH-fronting seems to be a preferred option, as for example in Evenki (Nedjalkov 1997: 7f.). For Even, on the other hand, WH-fronting is associated with emphatic/rhetorical questions; in regular constituent questions the interrogative pronoun remains in situ (Malchukov 2008). In Written Manchu, question words also remain in situ (Gorelova 2002: 222). In Udihe (Nikolaeva & Tolskaya 2001: 799), the position of focused elements including question words is strictly before the verb.

However, note that, according to the description by Girfanova (2002: 42), Udihe behaves like Evenki in putting the question word in sentence initial position. Even Nikolaeva & Tolskaya (2001: 799, 805) agree that the interrogative *ii-mi/j'e-mi* 'why' that is of converbal origin may optionally stand in clause initial position as well.

### 5.10.3 Interrogatives in Tungusic

Tungusic interrogatives have been treated in some detail before. The classical but partly outdated reconstruction can be found in Benzing (1956: 114f.). The most exhaustive lists of cognates that nevertheless lack many important data can be found in Cincius (1949: 264ff.) and Cincius (1975/77). Kazama (2003) elaborates on Cincius (1975/77) and also includes data from Kilen and Sibe but still is not exhaustive. Not to be underestimated are the data collected in Schmidt (1923a,b; 1928a,b) for Samagir, Samar, Ulcha, Nanai, Oroch, Udihe, Negidal, and Evenki. Of these, the first two varieties are almost unknown otherwise. Schmidt mentions Samagir *ekon* 'what' and Samar *xai* 'what', which is sufficient to classify the two as Ewenic (e.g., Evenki *ekun*) and Nanaic (e.g., Nanai *xai*), respectively (see also Doerfer 1978a). Table 5.114 gives an extended list of cognates for those five interrogatives that have the widest distribution among Tungusic languages. For the references, see the more detailed descriptions below. The use of Tungusic interrogatives or demonstratives as correlatives has recently been investigated in detail by Baek (2016: 185-226).

All languages except for some subdialects of Solon and Oroqen preserve the interrogative 'who'. The form has been reconstructed as *\*ŋüi* (Benzing 1956: 115) or *\*ŋui ~ \*ŋöi* (Kazama 2003: 68) for Proto-Tungusic and as *\*ŋii* for Proto-Ewenic (Janhunen 1991: 70f.). Only Kazama's reconstruction based on Ikegami is erroneous. The original *\*ü* regularly changed to *i* in Northern but to *u* in Southern Tungusic. In some Ewenic languages such as Solon or Oroqen as well as Udegheic, the velar nasal changed to an *n* while it apparently was lost in all of Jurchenic and Nanaic, except for Uilta and Ulcha. These are not regular developments but have certain parallels, e.g. Evenki *ŋina.kin*, Solon *nini.xin*,

Table 5.113: Question markers in Tungusic languages

Language	PQ	CQ	AQ
Even	=Ku#	-	2x =Ku#
Evenki	=Ku#	-	2x =Ku#
Evenki (Sakh.)	=Kuu#, =too#	-	2x =Ku#
Evenki (Kh.)	=Kv#	bei#	2x =li#
Negidal	Table 5.111, =Kʊʊ#, =i#	-, Table 5.111	Table 5.111, 2x =Kʊʊ#
Chaoyang Oroqen	=YEE#, =oo#, bAA#	-	2x =YEE#, 2x jOOmAA# (+2x =gUU)
Xunke Oroqen	=jA#, =ɔɔ#, bAA#	=jA#	(2x =jA +) aaki 'or', 2x jɔɔma#, 2x ɔɔmal#
Huihe Solon	=gi(i)#	(jeeme)	2x =gi(i)#
Ongkor Solon	=uu#, =ii#	-	2x =uu#
Udihe	=Ku#, =nA#	-	2x =Ku#, 2x =nA#, 1-2x -(e)s(i)#
Oroch	=nu#	-	2x =nu#
Nanai	=nu#	-	2x =nu#
Kilen	=nə#, =a#, =ma#, =ba#	-	2x =nə#, xəçi 'or' + =a#
Uilta	=(y)i#	-KA#	=(y)i#, yyuu#
Ulcha	=nʊʊ#	?=nʊʊ#	2x =nʊʊ#
Manchu	=o#, =n(i)#, =nio#, =nA#, =nu#	=o#, =n(i)#	Table 5.112
Aihui Manchu	=no#	-, ?-je#	xɛʂl 'or'
Yibuqi Manchu	=no#, mA#	-	xɛʂl 'or'
Sanjiazi Manchu	=no#, =nu#, ba#	=ni#, ?-jA#, =nə	haishi 'or'
Sibe	=na(a)#, ba#, ?-jə#	- ?-jə#	2x =na(a)#
Alchuka	=(k)ɔ#	=n(i)#, ?	?
Bala	=ɔ#, =ŋɔ#	=ɔ#, ?	?

Uilta *ŋinda*, Nanai *enda*, Manchu *inda.hūn*, but Udihe *in'e.i* ‘dog’ (cf. Benzing 1956: 68). The short vowel in some northern Tungusic languages must be a secondary innovation that is partly shared by the interrogative *i(i)*-. Kilen *ni* was borrowed from Udegheic, and Kili *ŋii* from Ewenic. A form *p'ə* ‘who’ mentioned by Mu Yejun (1986: 14) for Alchuka is most unexpected and cannot be explained with the reconstructed form *\*ŋüi*. Problematically, a [*pʰ*] in Alchuka usually corresponds to an *f* in Manchu (e.g., Alchuka *p'i*, Manchu *fi* ‘brush’) and Manchu *we* clearly corresponds to Nanai *ui* (e.g., Manchu *wesi-hun* and Nanai *uisi* ‘up’). It is not very plausible to assume that Nanai *ui* or Manchu *we* are not related to Uilta *ŋui* or Ulcha (*ŋ*)*ui*. Assuming that the Alchuka form is not a mistake, it is most likely related to Manchu *we*, but details remain obscure for the moment.

Table 5.114: List of cognates of five Tungusic interrogatives

Language	who	what, which	which, what	how many	how
Proto-Tungusic	*ŋüi	*ja-	*Kai	*Kadu	*Kooni
Even	ŋi	ja-k	i-rəə-k	adi	on
Arman	ŋii	jaa-ḵ	iää-ra-k	aadii	oon
Evenki	ŋi	e-kun	i-r	ady	oon
Negidal	nii, n'ii, ŋii	ee-xun, ee-kun	ii-	adii	oon(i)
Solon (Huihe)	nii, aawu	o-xon	ii	adi	iittü
Solon (Ongkor)	a(γ)uu	jo-xon	i(i)	adi	
Oroqen (Nanmu)	awu	i-hun	i-r(i)	adi	ooni
Oroqen (Chaoyang)	nii	r-kən	i-ri	?	ɔn
Kha. Evenki B.	nii	i-kun	ii-r	adii	oon
Kha. Evenki U.	nii	ie-kun	ii-r	adii	oon
Udihe	ni(i)	j'e-u	ii-	adi	ono
Oroch	n'ii	jaa-u	i-	ady	oni
Kili	ŋii	?ii- ~ e-	i-	adii	
Kilen (ZhY)	ni	ja	-	adi	ɔni
Kilen (ZhP)	ni	ja	-	ati	ɔməçi
Kilen (AJ)	ni	ja	χai	χadi, χadu	on(n)i
Kilen (Ling)	ui	-	hai	hadu	honi-bifi
Nanai	ui	-	xar	xado	xooni
Uilta	ŋui	-	xai	xasu	xooni
Ulcha	(ŋ)ui	-	xai	xadu	xoon(i)
Alchuka	?p'ə	?	kai-	kutu	katiram
Bala			a(i)-		
Manchu	we	ya	ai	udu	absi, adarame
Aihui Manchu	və	ja	ε	ɛdik	avçe
Sanjiazi Manchu	wə	ja	aj	udə	adəlmən
Yibuqi Manchu	və	jA	ei	utu	atər(ə)mə
Sibe	və	ya	ai	ut	afš(e)

Within the interrogative system of Tungusic *\*ŋüi* has a special position as it is unrelated to the other interrogatives. The same is also true for *\*ja-* ‘what, which’. Benzing’s (1956) Proto-Tungusic reconstruction *\*jaa-* and Janhunen’s (1991: 70f.) Proto-Ewenic reconstruction *\*ie-* seem to show the wrong vowel quantity and quality, respectively. As

for the development of the vowels, note a parallel development in Tungusic \**jaa-sa* ‘eyes’ > Evenki *ee.sa*, Borzya *ii.sa* etc. (Benzing 1956: 25; Janhunen 1991: 34). Only Nanaic languages have no reflex of \**ja-*, Kilen being a special case as the interrogative *ja* has been adopted from Udegheic or, less likely, from Jurchenic. Li Linjing (2011: 199) mentions a Kilen form *ya.o*, for which only Udihe *j’e.u* or perhaps Oroch *jaa.u* can be the source. The extension seen in this form exists only in northern Tungusic. Kili seems to have variation between *ii-* ~ *e-*, derived from Ewenic.

The interrogative \**Kai* (Benzing 1956 reconstructed \**xai*) is preserved in all branches but is absent in some parts of Kilen and exists only in relics in Udegheic. Benzing assumed the presence of a suffix attached to a stem \**xa-*, but no direct evidence for this has been found. The initial consonant has been regularly lost in most of northern Tungusic and Jurchenic and in most cases changed to a x-like sound in Nanaic. Kazama (2003: 56, 75) did not recognize the connection between *i(i)-* and Nanai *xai* etc. Admittedly, the stem extension (e.g., Evenki *i-r* ‘which’) can only be found in northern Tungusic. However, this must be a secondary innovation of some Ewenic languages that spread from the demonstratives (e.g., Evenki *e-r* ‘this’, *ta-r* ‘that’).

There is a certain amount of confusion surrounding the relation of the two stems \**ja-* and \**Kai-*. For instance, Doerfer (1985: 27) tried to show that they go back to one form, but his explanation is extremely speculative and does not appear to be actually based on any hard evidence. Nevertheless, the two forms are problematic as they have several properties in common, and are partly interchangeable. First, while northern Tungusic languages have an interrogative verb based on \**ja-*, the interrogative \**Kai-* has both nominal and verbal properties in Nanaic. This interesting difference can be shown with the help of Nanai and Kilen, which has been strongly influenced by Udihe in this regard (Table 5.115).

Table 5.115: Ambiguous interrogative stems in Nanai (Kazama 2007) and Kilen (Zhang 2013: 162)

	Nanai	Kilen	Meaning
verbal	<i>xai-xa-ni</i>	<i>ja-xəi-ni</i>	what- <small>PSR</small> -3 <small>SG</small>
nominal	<i>xai-wa</i>	<i>ja-wə</i>	what- <small>ACC</small>

Ewenic and Udegheic roughly pattern with Kilen while Jurchenic is close to Nanai but apparently is unique in showing an obligatory verbalizer (e.g., Manchu *ai-na-*, Bala *a-na-*, Alchuka *kai-na-*). However, Udihe also has an optional derived form *j’e-ne-*. This split has not only been overlooked by Benzing (1956) but also by several other scholars such as Tolskaya & Tolskaya (2008: 99). Interestingly, some forms have the same derivation but are based on different stems. For example, interrogatives meaning ‘why’ usually have a verbal basis and are really converb forms of the interrogative verb, e.g. Even *ja-mi*, Udihe *j’e-mi*, but Nanai *xai-mi* and Manchu *ai-na-me*. Apart from Oroch, Udihe is exceptional as it also has the form *ii-mi* that can be compared with Nanai. Second, the two interrogative stems are partly interchangeably in Jurchenic. In Manchu dialects, for example, there is

synchronic variation between alternative forms such as *ai-erin* ~ *ya-erin* ‘when’ or *ai-ba* ~ *ya-ba* ‘where’ without apparent differences in meaning (see below for more examples).

However, the fact that languages as distantly related and located as Even and Manchu have traces of both stems is clear evidence for their existence at a very early stage in the development of Tungusic. Furthermore, *\*Kai* is part of a larger group of interrogatives that share a resonance in *\*K-* that most likely is etymologically connected. But even in Proto-Tungusic their exact derivation must have already been obscure. For example, Benzing, based on the assumption of analyzability of *\*xa-i*, reconstructed the interrogative meaning ‘how’ as *\*xaoni*, which has to be rejected, as there is no indication of an original diphthong. Many modern languages preserve a long vowel, which is why I reconstruct the form as *\*Kooni* instead. Janhunen (1991: 70f.) assumed a stem *\*xoo-*, but there is no clear evidence that *-ni* might have been a suffix. This interrogative is preserved everywhere except for Solon and Jurchenic. In a similar vein, Benzing’s (1956) reconstruction *\*xaduu* ‘how much’ with a long vowel has no real basis as most languages simply have a short vowel. Based on the distribution of northern Tungusic *i* and southern Tungusic *u*, but Nanai *o*, as well as a comparison with Mongolic (on which see below), the form may probably be reconstructed as *\*Kadu* instead. In the latter two interrogatives there are some irregular developments such as a progressive vowel assimilation in Udihe *ono* ‘how’ (cf. Oroch *oni*) and a retrogressive assimilation in Jurchenic, e.g. Manchu *udu* ‘how many’ (cf. Ulcha *xadu*).

Benzing assumed a Proto-Tungusic resonance in *\*x-*. But in my opinion, new evidence (e.g., Mu Yejun 1986; Hölzl 2017b) points to a possible reconstruction as plosive (see also Rozycki 1993). This assumption is based on data from Alchuka that exhibit what could be a conservative feature lost in all other Tungusic languages. However, given its unclear phonetic status, for now I use a label *K-* in the reconstructions instead. The limited data from Alchuka contain four interrogatives with an initial unaspirated velar plosive *k* (or perhaps *g*) that is not present in Manchu (Table 5.116). It has been suggested to me by András Róna-Tas (p.c. 2015) that the consonant might be a secondary innovation in Alchuka. The initial consonant is attested in about two dozen instances, and it may well be a secondary innovation in some of them. However, the fact that it systematically appears in many attested interrogatives and has a correspondence in Nanaic *x-* suggests that at least in this position it should be of Proto-Tungusic origin.<sup>31</sup>

From the typological criterion adopted in this study, interrogatives in Alchuka qualify as *K-*interrogatives. Regardless of the exact reconstruction that I intend to clarify in future studies, Proto-Tungusic clearly has to be classified in the same way.

Benzing (1956: 114f.) has three more reconstructions (*\*xalii* ‘when’, *\*xason* ‘how much’, and *\*xaba-sikii* ‘whither’), all of which exhibit several deficits. Only the last one is attested in Jurchenic languages. The first two may perhaps be corrected to *\*Kaali* and *\*Kasu(n)* (see the description of individual languages below). The last form *\*xaba-sikii* poses several problems that cannot be solved easily, but I propose the slightly different reconstruction *\*Ka-bV-si-ki(i)* instead (Table 5.117). Apparently we are dealing with a case form, more exactly a directive, of an otherwise unknown interrogative starting with *K-*

<sup>31</sup>I am currently preparing a more detailed investigation of the problem.

Table 5.116: Selected Alchuka interrogatives (Mu Yejun 1985; 1986; 1987; 1988b,a) with Manchu cognates (Norman 2013); inner-Tungusic loanwords are in parentheses

	Alchuka	Manchu
for what reason	( <sub>_ei</sub> ) t'uku	<sub>_ai</sub> turgun
how	katiram	<sub>_adame</sub>
how many	kutu	<sub>_udu</sub>
to do what	kai-na-mei	<sub>_ai-na-mbi</sub>
what	( <sub>_ei</sub> )	<sub>_ai</sub>
what has happened	gai-na-hanbie	<sub>_ai-na-habi</sub>
what (is it)	kent'aka	<sub>_antaka</sub>
when	( <sub>_ant'</sub> an̄gi)	<sub>_atanggi</sub>
why	( <sub>_einu</sub> )	<sub>_ainu</sub>

Table 5.117: Cognates of \**Ka-bV-si-kii*

Language	Form	Meaning
Even	awaskii ~ awuskii	whither
Negidal	awaskii	whither
Evenki	awaskii	whither
Oroch (Schmidt)	avasee	whither
Kilen (An Jun)	χaoçi	whither
Ulcha (Schmidt)	xavasi	whither
Nanai	xaosi	whither
Manchu	absi	whither, how
Sibe (Zikmundová)	afš(e)	whither, how

that has parallels in the demonstratives, e.g. Even *ə-wə-ski(i)* ‘in this direction, hither’, *ta-wa-ski(i)* ~ *ta-wu-ski(i)* ‘in that direction, thither’, and *a-wa-ski(i)* ~ *a-wu-ski(i)* ‘in what direction, whither’ (Benzing 1955: 77f.; Benzing 1955: 86, 113f.). Manchu preserves the forms *ebsi* ‘hither’ (Alchuka *kə'uzi*), *yabsi* ‘how very’, and *absi* ‘whither’ which has acquired the meaning ‘how, why’.

Sibe *afš(e)* is a regular continuation of Manchu *absi*. It seems possible that the final element was only present in Ewenic but not in the other Tungusic languages. Note that there are several case forms that may either stand alone or may be combined with a comparable suffix, e.g. Solon dative *-dU*, ablative *-dU-xi*. Apart from the case suffix, there is another element \**-bV*, possibly of nominal origin, that might also be present in Proto-Tungusic \**Ka-bV-gu(u)* ‘which’, an interrogative that had not been reconstructed by Benzing (1956) (Table 5.118). Even *aw-gic* ‘whither’ could go back to the same source \**Ka-bV-*

5 Survey of the grammars of questions in Northeast Asia

Table 5.118: Cognates of \**Ka-bV-gu(u)* ‘which one > who’ (cf. Hölzl 2014b); Schmidt = Schmidt (1923b), Grube = Grube (1900), Castrén = Castrén (1856)

Language	Form	Meaning
Uilta	xaawu	which (one)
Ulcha (Schmidt)	xavu	which (one)
?Nanai (Grube)	xawui	which (one)
Evenki (Castrén)	awguu ~ abguu	which (one)
Evenki	avgu	which (one)
Evenki (Khamnigan)	abguu	which (one)
Even	awug ~ awag	which (one)
Negidal	L avvu, avgu, au, U avgavu	which (one)
Solon (Hailar)	ayuu	who
Solon (Huihe)	awu	who
Oroqen (Nanmu)	awu	who

In Uilta, the initial consonant has been preserved as *x-*, but intervocalic \**VbV* and \**VgV* have both been regularly lost (Benzing 1956: 30, 34). The final \**uu* must have changed to *wu* following the newly formed long vowel *aa*. Uilta in addition has a special accusative form *xaakkoo* (Tsumagari 2009b: 4, 7f.) which might indicate the presence of an earlier consonant other than *w* since only stems ending in -CV show this type of assimilation of the accusative marker -*BA* and the geminate *kk* indicates a plosive. This consonant may have been a relic of the original \**g*. My reconstruction is almost identical to Kazama’s (2003: 68) \**xabagu*. But the vowel in the second syllable is not entirely certain as it has been lost in several languages and shows variation between *a ~ u* in Even. The intervocalic \**VbV* changed to *w* in northern Tungusic languages. The Even variant *awu.n* indicates that the final \**-gu(u)* is a suffix that replaces the unstable nasal. In Solon, the \**b* > *w* was lost and the \**g* changed to *γ*. After the *γ* had been lost in some Solon dialects, the final long vowel must have changed to *wu* as in Uilta. The second possibility that Solon *ayuu ~ awu* goes back to a form without the suffix \**-gu(u)* in which the \**b* changed to *γ* is less likely due to the presence of a long vowel that can only be traced back to the suffix. Oroqen *awu* is a Solon loanword. Some points remain unclear, however. For example, does Khamnigan have a *b* instead of the expected *w*, because of the following *g* and how does the Upper Amgun Negidal form *avgavu* fit into the picture? Possibly there was a variance between different suffixes, such as in Evenki *idy-vu, idy-gu* ‘which one’. Cincius (1975/77: 4f.) includes Manchu *absi* ‘how’ in the list of cognates, which is clearly a mistake.

There is one rather problematic interrogative that has several functions and can have both verbal and nominal properties. In interrogative sentences the meaning is extremely broad as it may be translated as ‘who’, ‘what’, ‘which’, ‘where’ or ‘how many’ (Bulatova & Grenoble 1999: 24). Given its unclear semantics it has been glossed as INT (interrogative). Consider the following example from Evenki.

(375) Evenki

*aŋii aŋii-βa aŋii-ẓ̌a-ra-n?*

INT INT-ACC INT-IPFV-PRS-3SG

‘Who is doing what?’ (Bulatova &amp; Grenoble 1999: 26)

Problematically, the word may also be used in declarative sentences where it may “replace nearly any verb” (Bulatova & Grenoble 1999: 26) or may also function as a demonstrative. Given that cognates from the Nanaic branch do not show an initial consonant, this word is clearly of a different origin than the other interrogatives. The best treatment of this unusual word has been given by Idiatov (2007: 301ff.). Elaborating on Cincius (1975/77), he gives the following account. The word started out as a noun meaning something like ‘thing’, which in Evenki may have been combined with the genitive or the alienable possession marker. The second step was the development of a “placeholder or filler”, such as English *whatchamacallit* (Idiatov 2007: 302). This function is attested in several other Tungusic languages. The last step was from a placeholder to an interrogative. Since the last function is restricted to Evenki, the forms from other languages will not be treated here any further.

Tungusic interrogatives exhibit several striking similarities to Mongolic that cannot be explained by chance (Table 5.119). These comparisons do not stand on their own but join well-known similarities in the personal pronouns and demonstratives.

Table 5.119: Similar interrogatives in Mongolic and Tungusic

	Mongolic	Tungusic
what	*ya-xu/n	*ja- (northern + *-ku/n)
to do what	*ya-xa-	*ja- (only northern and Kilen)
how many/much	*kedü-	*Kadu
when	*keli	*Kaali

This is not the place to present a discussion of a possible genetic connection between Mongolic and Tungusic, but it should be pointed out that language contact could also account for these similarities (see §5.8.3). Most likely, the forms have been borrowed by Tungusic because the morphology involved is also known from other elements in Mongolic, such as demonstratives, e.g. \**e.li*, \**te.li* or \**e.dü-*, \**te.dü-* (Janhunen 2003d: 20). The first two from the list also appear to have been borrowed by Nivkh (§5.2.3)

The following will address interrogatives in the individual branches of Tungusic in turn. Table 5.120 gives some interrogatives from Arman and Even. Even has some unique developments in interrogative paradigms (Table 5.121). While the stem extension of the interrogative *i-*, an extension from the demonstratives, is shared by most Ewenic languages, Even *i-rəə-k* exhibits a further innovation. The final *-k* stems from the interrogative *ja-k* and even found its way into the demonstratives and may tentatively be analyzed as a newly formed nominative marker that is restricted to these four stems.

5 Survey of the grammars of questions in Northeast Asia

Table 5.120: Interrogatives in different dialects of Arman and Even (Doerfer & Knüppel 2013, modified; Benzing 1955; Sotavalta 1978: 12, passim, modified; Schiefner 1874); DK = Doerfer & Knüppel, B = Benzing, S = Sotavalta, Sch = Schiefner; case forms and several alternatives are not shown

	Arman (DK)	Even (B)	Western Even (S)	Anadyr (Sch)
who	ɲii, ɲii	ɲi	ɲii	ɲi
what	jaa- <b>k</b>	ja- <b>k</b>	je- <b>k</b>	ja- <b>k</b>
which	iää- <b>ra-k</b>	i-rəə- <b>k</b>	i-rä- <b>k</b>	i-
to do what	jaa-	ja-	ie-	(i)jä-
why (CVB)	jaa-mi, jaa-ɲgai	ja-mi	ie-gaji	(i)jä-m(i)
which one		awug		
where	i-lee, iää- <b>laa</b>	i-ləə	i-lä, i-la, i-DDu	i-la
how	oon, uun	on	<sup>i</sup> etat, <sup>i</sup> ekat	on
when	ook	ok		ok
how much/many	aadii, adal	adi	adi	
how much/big	ašun	asun	ahun ~ ahuun	

Table 5.121: Nominative and accusative case forms of interrogatives and demonstratives in Even (Benzing 1955: 77, 79)

	who	what	which	this	that
NOM	ɲi	ja- <b>k</b>	i-rəə- <b>k</b>	ə-rəə- <b>k</b>	ta- <b>ra-k</b>
ACC	ɲi-w	ja-w	i-rə-w	ə-rə-w	ta- <b>ra-w</b>

The Evenki and Negidal interrogative systems are extremely similar to one another. Generally, the forms tend to be a bit longer than those in Even. In Khamnigan Evenki, while the Urulyungui dialect preserves a small difference between *ie-* and *ii-*, the two interrogatives \**ja-* and \**Kai-* completely coalesced into *i(i)-* in the Borzya dialect. In general, Khamnigan Evenki interrogatives appear to be more closely related to Oroqen than to Evenki. Apart from this partly shared sound change, both groups have also changed the initial velar nasal to an alveolar nasal in *nii* ‘who’ and have a form *aali* ‘when’ instead of *ookin* in Evenki. But apart from *awu* ‘who’, which is borrowed from Solon, Oroqen does not have a cognate of *abguu* ‘which one’. The Khamnigan form *iir-giiji* ~ *iir-giid* has a cognate in Oroqen *iri-gidə* and Evenki *ir-git*. Evenki dialects, such as the one from Sakhalin, exhibit a very similar interrogative system but shows some regular phonological differences (e.g., *axun* ‘how many’, Bulatova & Cotrozzi 2004; Atknine 1997).

There are descriptions for several Oroqen dialects, the interrogatives of which are given in Table 5.123. Only a selection of case forms is included. The complex forms *ixun-tšaalın* or *ikvn džaalin* ‘why’ and *adi erin-du* ‘when’ contain the Manchu loanwords *jalin* ‘reason’ and *erin* ‘time’. The second part in *iri-gətfin* ‘what kind of’ is probably not

Table 5.122: Evenki (Nedjalkov 1997: 3-18, 135-136, 214-216, 318ff.), Negidal (Cincius 1982: 34, *passim*), Khamnigan Evenki (Janhunen 1991: 70f.), and Aoluguya Evenki interrogatives (Hasibate'er 2016: 171, 238); U = Upper Amgun, L = Lower Amgun dialect of Negidal. B = Borzya, U = Urulyungui dialect of Khamnigan Evenki

Meaning	Evenki	Negidal	Khamnigan	Aoluguya
who	ŋi	nii, n'ii, ŋii	nii	ni:, nɪ:
what	e-kun	L ee-xun, U ee-kun	B i-kun, U ie-kun	æ:-qɔŋ, æ:-ʁɔŋ
what kind of (-dy 'ADJ')	e(e)-ku-dy			
what kind of (PL) (-ma 'ADJ')	e(e)-ku-ma			
why (-da 'CVB')	ee-da	ee-daa(j)		
how much	ady	adii	adii	adɪ, addi
how many times	adyra	adiijaa		
by what number (-tal 'QUANT')	ady-tal	adiital		
which	anty			
what, which	aŋi	aŋ, aŋi, aŋe 'yes, alright, okay'		
how many which, what kind	asun avady	asun		irəgeçin, irgə:tʃin
whither		awaskii		
what, which	avgu	L avvu, avgu, au, U avgavu	abguu	
which	i-r	ii		ɪ:-ra, i:ʂ
where	i-du	ii-duu, ii-laa	ii-dvɪ, B i-lee, U ii-lee	i:-du, ɪ:-la, ɪ:-ra
which one how much/ many	idyvu, idygu oki		ooki	ɪ:rba, ɪrɕa:
when	ookin	L ooxin, U ookin	aali	ɔ:qin, ɔʁ:in
how	oon	oon(i)	oon	æ:χa

5 Survey of the grammars of questions in Northeast Asia

Manchu *hacin* ‘kind, sort, class, item’ (from Korean) because there is a similar suffix in other Ewenic languages (Benzing 1956: 100), e.g. Evenki *-gAchin* ‘similar to, just as, like’ (Nedjalkov 1997: 56), e.g. Aoluguya Evenki *irəgecin ~ irgə:tfin*. The suffix *-du* is a locative and dative case marker that can also be found in *oki-du* ‘when’ (based on *oki* ‘how much’, influenced by Solon), *ixu-tu* ‘when’ (based on *i-xun* ‘what’ with stem extension) and *(i)i-tu* ‘where’ (based on *i-(xun)* ‘what’ without stem extension). The etymology of *idzirgee* ‘which one’ remains unclear for now. The Nanmu forms *awu* ‘who’, *oonde* ‘what kind of’, and *joonde* ‘why’ have been adopted from Solon. The same is probably true for *iktu* ‘how’ as mentioned by Chaoke. The origin of *jEE-ma* ‘which one’ remains unclear, but a connection to Mongolic seems plausible. One can observe a slow phonological convergence of the two different stems *i-hun* ‘what’ and *i-r(i)* ‘which’.

Table 5.123: Interrogatives in different dialects of Oroqen (Hu Zengyi 2001: 101, 148, 261; 1986: 94; Chaoke D. O. 2007: 47, 257f.; Zhang Yanchang, Li Bing, et al. 1989: 56, 141, 184; Han Youfeng & Meng Shuxian 1993: 43, 264-265; CK = Chaoke D. O. 2014a: 164-165, passim)

	Chaoyang	Wulubutie	Nanmu	Xunke	Shengli	CK
who	nii	nii	awu	nii	nii	ni
what	ikɔn	ikɔn	ihun	ikɔn	ixun	ikun
when	aali(t)	aali	aala	aali	aali, <b>ixu-tu</b>	oki-du, <b>adi</b> <b>erin-du</b>
how many			adi		ati	adi
what kind of	iri-gətʃin, iŋŋətʃin, in- nuwəən	iri-gətʃin, iŋŋətʃin, in- nuwəən	<b>oondi</b>	iri-gətʃin	iri- kətʃin, iŋNetʃin	
which	i-ri	i-ri	i-r(i)		(i)i-ri	i-ri
which one				<b>jEEma</b>	<b>jeman</b>	
which one (PL?)	idzirgee	idzirgee		idzirqə	itʃərkee	
where	i-ləə, i-du		i-lə	i-rə	i-ləə, iitu	i-le(-ni/ ŋi)
whither	iri-gidə	iri-gidə	iri-gidə	iri-gidə	iri-kita, iir-tə	
why	ɪda, <b>ikɔn</b> <b>dʒaalin</b>		iida, iima, <b>joonde</b>		<b>ixun-</b> <b>tʃaalin</b>	
how	ɔɔn	ɔɔn	ooni	ɔɔn	(w)ɔɔn	oon, <b>iktu</b>
how much	ɔɔki		oohi	ɔɔki	ɔɔxi	(o)oki

Solon interrogatives are probably the most aberrant among Ewenic languages. The interrogative *ni(i)* ‘who’ has been almost completely replaced (see Table 5.124). The unexpected vowel quality in *(j)o-xon* ‘what’ can possibly be attributed to influence from Dagur (*yoo(n)* ‘what’). *ohi-du* ‘when’ similar to Evenki is based on *ohi* ‘how much’, but contains an additional locative marker. This form has been adopted by one Oroqen di-

Table 5.124: Interrogatives in different dialects of Solon (Chaoke D. O. 2009: 35f., 250ff., 351f., 355; Tsumagari 2009a; Poppe 1931: 110); CK = Chaoke, T = Tsumagari, P = Poppe, R = Ramstedt (Aalto 1976; 1977, modified), K = Kamimaki (Lie Lie 1978: 175, 177, modified); case forms are not listed

	Huihe (CK)	Huihe (T)	Hailar (P)	Ongkor (R)	Ongkor (K)
who	awu, <b>ni</b>	aawu, <b>nii</b>	ayuu	auu ~ ayuu	ahu
what	o-hoŋ	ii, o-xon	ii	i(i), jo-xon	o-hon, u-hun
when	ohi-du	ooxi-du	ooxii-du	<b>aali</b>	oke-du
what kind of	ondi	oondii	oondii	oondin	onde
which	iggʉ	<b>ii-r</b> , iggü	iixuuu		
where	i-lə	ii-lee	ii-ləə	ii-lə	i-le, i-lo
how	ittʉ	iittü	iittuuu		
why	ida, joodoŋ	yoodon	iidaa, joodaa	ida, joxon-du	yoda
how many	adi	adi	adii	adi	aade
how much	ohi	ooxi	ooxii, <b>oondii</b>	ookr	ooke, <b>one</b>

alect while Ongkor Solon *aali* ‘when’ in turn can perhaps be traced back to influence from Oroqen. Dagur *yoondaa* ‘how, why’ is perhaps the source of *joodaa* ‘why’.

The origin of *iggʉ* ‘which’ and *ittʉ* ‘how’ is unclear, but in Solon a geminate suggests the earlier presence of a consonant cluster as can be seen in many examples, e.g. Evenki *irgi*, Solon *iggi* ‘tail’. Possibly, the forms are based on the stem *i(i)-r(i)*, followed by a case ending. At least synchronically the form with the suffix *-r(i)* has no wide distribution among Solon dialects, which usually employ the bare stem *i(i)*. But further evidence for this view can be gleaned from the demonstratives *e-ri* ‘this’ and *ta-ri* ‘that’ that still have the extensions, and the derived forms *ettü* ‘in this way’ and *tattü* ‘in that way’ (Tsumagari 2009a: 3, 6). Problematically, from a synchronic perspective no case marker has the expected form *\*-gü* or *\*-tü*. At least the latter may have a connection with the dative *-du* ~ *-dü* that in Evenki also has a variant *-tu* with an unvoiced consonant. Ramstedt’s Ongkor Solon materials have been recorded in Tacheng (Lie 1978: 128). From the city Alimtu, Muromskij collected several unproblematic forms including *au* ‘who’, *ad’* ~ *adi* ‘how many’, *ile* ‘whither’, *ida* ‘why’, *on(i)* ‘how’ (Lie 1978: passim, Kałużnyński 1971: passim). Chaoke D. O. et al. (2014: 63) mention a form *antie* ‘how’ that seems to correspond to Evenki *anty* ‘which’.

*Kili* is a mixture of Nanaic with Ewenic elements (Doerfer 1978a), but judging from the interrogatives alone, *Kili* appears to have stronger affinities to Ewenic than to Nanaic (e.g., *ŋii* ‘who’, *e-ma* ‘what’, *adii* ‘how many’, *ali* ‘when’, *i-du* ‘where’, *ii-daj* ‘why’, *osi* ‘which one’, Kazama 2003: passim; see Sunik 1958 for details). Note the characteristic

5 Survey of the grammars of questions in Northeast Asia

form *ɲii* as well as the absence of the initial consonant *x-*. Kili *osi* remains obscure but has a cognate in Kilen *ɔci*. The two interrogatives *iidaj* and *ema* are also characteristic of Ewenic, but might stem from two different sources as indicated by the different length and quality of the initial vowel.

As expected, interrogatives in Udegheic languages show affinities with Ewenic. Table 5.125 gives an overview of some forms attested in Udihe and Oroch.

Table 5.125: Udihe (Nikolaeva & Tolskaya 2001: 348ff.; Tolskaya & Tolskaya 2008: 100) and Oroch interrogatives (Avrorin & Boldyrev 1978; Lopatin 1957, collected in 1924, modified); not all variants are listed

	Udihe	Oroch	
	NT	AB	Lopatin
who	ni(i)	n'i	n <sup>(v)</sup> i
what	j'e-u	jā(-ʊ)	ya(-u)
what kind of	je-me		
why	je(-ne)-mi, ii-mi	jæ-mi, jǎ-mi, jā-na-mi, jī-mi	
what, where, why	j'efe	jāvʊ	yava
when	ali	āli	ale
how many	adi	adi	ady
how	ono	ōn'i	oni
which	onobui	ōn'i bi	

Nikolaeva & Tolskaya (2001: 348) claim that *j'e-fe* 'what, where, on which place' is an accusative form. It seems, however, that this form rather corresponds to Manchu *ya-ba* and Sibe *ya-va* 'where, which place'. The strange looking form *onobui* is probably a contraction of *ono* 'how' with an inflected form of the copula *bi-*, which has a parallel in Kilen. Apart from this, several more Udihe interrogatives have been adopted by Kilen as well (see below).

The interrogative *ni(i)* 'who' is declined as the word *nii* (or *ninta*) 'man' but does not have an etymological connection to it, as claimed by Schulze (2007). Instead, the forms correspond to Nanai *ui* and *nai*, respectively, and are similar only by chance. But one cannot exclude the possibility of a folk etymological connection. The interrogative *j'eu* exhibits some irregularities. Apart from the nominative forms, the paradigms are parallel to the demonstratives (Table 5.126). The ending *-u* in *j'e-u* or *jaa-u* 'what' is identical in origin with Evenki *-kun* in *e-kun*. In Oroch, but not in Udihe, this extension is also found in most case forms. This is a secondary leveling that has a parallel in Evenki and Oroqen. Udihe *j'euxi* 'whither' probably corresponds to Nanai *xaosi* but is based on a different stem. As in Nanai the same suffix *-uxi* can otherwise only be found in the demonstratives.

In Udihe, forms such as the dative *j'e-du* or the locative *j'e-le* have the variants *ii-du* and *ii-le*. According to Nikolaeva & Tolskaya (2001: 349), this only represents a difference in pronunciation. However, *ii-* really is the relic of a different stem of which no nominative

Table 5.126: Interrogative and demonstrative paradigms in Udihe (Nikolaeva &amp; Tolskaya 2001: 100, 343f., 348) and Oroch (Avrorin &amp; Boldyrev 2001: 193, 197)

	Udihe				Oroch			
	who	what	this	that	who	what	this	that
NOM	ni	j'e-u	ee, ei	ute, uti, tee, tii	n'ii	jaa-u	ei	tei, ti
ACC	ni-we	j'e-we	a-wa	(u)ta-wa	n'ii-ve	jaa-va	ee-ve	taa-va
DAT	ni-du	j'e-du	o-du	(u)ta-du	n'ii-du	jaa-u-du	ee-du	taa-du
LAT	ni-tigi	j'e-uxi	a-uxi, i-tigi	(u)ta-uxi, (u)ta-tigi	n'ii-jiji	jaa-u-jiji	ee-jiji	taa-jiji
LOC	ni-le	j'e-le	o-lo	(u)ta-la	n'ii-le	jaa-u-la	ee-le	taa-la
PROL	ni-li	j'e-li	o-li	(u)ta-li	n'ii-li	jaa-u-li	ee-li	taa-li
ABL	ni-digi	je-digi	o-digi	(u)ta-zi	n'ii-dui	jaa-u-dui	ee-dui	taa-dui
INST	ni-zi	j'e-zi	o-zi	(u)ta-digi	n'ii-ji	jaa-u-ji	ee-ji	taa-ji

Table 5.127: Selected case forms of two different interrogatives in Udihe, Even, and Manchu

	Udihe		Even		Manchu	
NOM	j'e-u	-	ja-k	i-rəə-k	ya	ai
DAT	j'e-du	ii-du	ja-du	i-du	ya-de	ai-de
LOC	j'e-le	ii-le	ja-la	i-ləə	-	-
CVB	j'e(-ne)-mi	ii-mi	ja-mi	-	-	ai-na-me

or citation form is left in Udihe (Table 5.127). Oroch likewise has these alternative forms, e.g. *i-du* (Schmidt 1928a).

In Manchu the locative *\*-la* is only preserved in relics (e.g. *ama-la* 'behind') and the stem extension is restricted to the demonstratives *e-re* 'this' and *te-re* 'that'. Strangely, Udihe also shows this variation between two stems in the interrogative *ii-mi* ~ *j'e-mi* 'why'. Given that these are converb forms, Udegheic is the only branch in which both stems can function as interrogative verbs. Udihe *ii-mi* directly compares with Nanaï *xar-mi* and *j'e-mi* with Even *ja-mi*. Udihe furthermore has the variant *j'e-ne-mi*, which is similar to Manchu *ai-na-me*, but is based on the other stem.

Unlike all northern Tungusic languages and most of Jurchenic, Nanaïc interrogatives form a coherent system in which all forms share the resonance *x-*. The only exceptions is the interrogative meaning 'who' that was already different in Proto-Tungusic, as well as some isolated Uilta forms that have perhaps been borrowed from Nivkh (*sado*, *saa*, *nuulu*) (see §5.2.3). Patryk Czerwinski (p.c. 2018) elicited the forms *sadu* and emphatic *sadoo* from a northern Uilta speaker.

5 Survey of the grammars of questions in Northeast Asia

Table 5.128: Interrogatives in Najkhin Nanai (Kazama 2007), Ussuri Nanai (Sem 1976), Ulcha (Majewicz 2011), and Uilta (Ikegami (1997); Tsumagari (2009b); Majewicz (2011)); accents removed

	N. Nanai	U. Nanai	Ulcha	Uilta (M)	Uilta (TI)
who	ui	ui	nui	ŋuj	ŋui
what	xai	χay	xaj	xaj	xai
to do what	xai-	χay(-ra)-	xaj-		xai-
why	xai-mi	<b>χajɔ</b>			xai-mi
when	xaali	χāl'y	xali	xali	xaali
which one	xamačaa	χamatsa	xamata		<b>xaawu</b>
what for	xaigoji	χaonk'y, χaonk'y, χaonk'yd'y	xajd'u	xajbu	
where	xai-do	χay-dɔ		xaj-du, <b>sado</b>	xai-du
whither	xaosi	χaos'(y), χaɔs'y, χay-χaydu	xajban	xavasaj, xoty, <b>nuulu</b>	xawasai
whence	xajaǰi		xajdani	xavedu	xamačču
how	xooni	χōn'i	xon(i)	xooni	xooni
how many	xado	χado, χadɔ	xasu, xadam	xasu	xasu

With respect to the other Nanaic languages, **Kilen** exhibits a very different set of interrogatives (Table 5.129). Not only is there a rather confusing variation in the origin of the individual forms, but different accounts show striking differences as well. The forms that most closely resemble Nanai have been collected by Ling Chunsheng (1934) and they might actually represent Hezhen instead of Kilen. All other descriptions show variations between some forms of Nanaic and some of Udegheic or Jurchenic origin that have been borrowed. *χadi*, if this is not a typo, is an especially interesting form as it combines features typical of southern and northern Tungusic. It shares the initial consonant typical for Nanaic, but has a final *-i* that can only be of northern Tungusic origin. Udihe loans include *ni*, *adi*, *oni*, and maybe several more such as *ja*, *uki*, and *ɔnibici*, although the latter has also been recorded with an initial consonant atypical of Udihe. Manchu elements include the nouns *jaka* ‘thing’ (in *ia-mə-dzaka*) and perhaps *erin* ‘time’ (in *ia-ma-erin*, *adi/ya erin-du*, and *iarin*). The interrogative *ja* is certainly of Udihe origin, because Li Linjing (2011: 199) mentions a form *ya-o* ‘what’ that can only stem from Udihe *j'e-u* but not Manchu *ya*. The interrogative *iətin* ‘when’ appears to be a combination of *ja* and perhaps an otherwise unknown noun meaning ‘time’ or suffix that can also be found in Manchu *atanggi* ‘when’, see below. The forms *onnomi* and *onaqami* are obscure but may contain the converb marker *-mi*. As seen before, NDSSLD (1958: 82) mentions the

Table 5.129: Interrogatives in different descriptions of Kilen (Ling Chunsheng 1934: 243, 245; An Jun 1986: 38, 63; Zhang Yanchang, Li Bing, et al. 1989: 40, 44f., 70, 74f., 88, 144; Zhang 2013: 95, 162f.; Chaoko D. O. 2014b: 164 et passim); the table also contains all available case forms. L = Ling Chunsheng, AJ = An Jun, ZZD = Zhang et al., Z = Zhang, CK = Chaoko; forms from An Jun (1984) in square brackets

	L	AJ	ZZD	Z	CK
who	ui	ni	ni	ni	ni
what, which		ja	ia	ia/ja	ya
what	hai	χai, [χajə]			hay
what (thing)			ia-mə- dzaka		ya-ma
which		[ɔçə]	ɔçimkən	ɔçi	iri
where			ja/ia-du	ia-tu	i-lə, i-du
where from				ia.tu-tiki	
where to		χaoçi	ia-lə	ia-tu-lə	
when, what time			ia-ma-ərin	iarin	iətin, ya erin-du, adi erin-(du/ ni)
how	hənbifi	ombiçə, on(n)i, onnomi, [onbiçə]	ɔni, ɔnibiçi 'what'	ɔməçi	oni, ikti
how many	hadu	χadi ~ χadu, [adi]	adi	ati	adi
how much				uki	uki
why		onaqami	unakəmi	ɔɲnəmi, ɔməmi	

two Kilen forms *ya-le* 鴨勒 and *ali* 阿里.<sup>32</sup> Kilen *ɔimkən* is a contraction of *ɔci* and the numeral *əmkan* 'one', which is most likely of Jurchenic origin (Manchu *emken*). The interrogative *ɔci* has a cognate in Kili *osi* and *ikti* 'how' perhaps in Oroqen *iktu* 'how'. Both remain unclear etymologically.

Interrogatives in **Jurchenic** show marked differences from the other Tungusic languages. Almost no information is available for Bala and the few interrogatives available for Alchuka have already been given in Table 5.116. Sibe and Written Manchu, on the other hand, are exceptionally well described. Table 5.130 gives an overview of interrogatives in Manchuic languages. Aihui Manchu *edzəxə* ~ *edzəy*, *edik*, according to Enhebatu

<sup>32</sup>The character 里 was incorrectly written as 黑.

(1995: 149), has a cognate in Sanjiazi Manchu *aizug*, *aizig*, *aizvu* ‘how much, many’, but remains obscure. Sibe *yask(ə)* might be comparable as well, but seems to be based on *ya* instead of *ai*. The Manchu form *adarame* (Alchuka *katiram*) has never been analyzed in a clear manner. There are several possibilities, but the most likely scenario is a derivation from the interrogative *ai* that subsequently lost the *i* as in other forms. If the final *-me* is the imperfective converb form as a comparison with Nanai *xai-mi* ‘why’ might suggest, then at least one of the other elements present might have been a verbalization. Both *-da* and *-ra* are attested in this usage, but their combination would be most unusual. Problematically, the verbal interrogative in Manchu has the regular form *ai-na-* (Bala *a-na-*, Alchuka *kai-na-*). In fact, Manchu also has the expected form *ai-na-me* ‘how’. Perhaps the form has to be analyzed as *\*a(i)-da-ra-me* with an unclear derivation of the interrogative stem. The forms *ainu* and *antaka* (Alchuka *kent’aka*) are even more obscure but probably derive from *\*Kai-*, too.

Unfortunately, there is almost no record of **Jurchen A** interrogatives. However, there is a form that has been reconstructed as *\*wanon* 晚灣 ‘how’. Kiyose (1977: 137) hypothesized that it might be connected to Manchu *antaka*, but this is not very convincing. At first glance, no similar looking form is attested in any other Jurchenic language. The only tentative solution that I can think of, apart from an altogether unknown interrogative or a mistake, is to compare *\*wanon* with Manchu *ai-na-me*, which has the same meaning and shows some remote formal resemblance. The initial *\*w-* is extremely problematic but could perhaps be a reflex of an initial consonant, cf. Alchuka *kai-na-*. The interrogative *ai* lost the *i* in some other instances as well, cf. Bala *a-na-*. What has been reconstructed as *\*-n* might thus be a converb form that does not, however, match Manchu *-me* or Bala *-mi* (Mu Yejun 1987: 30). The converb has been recorded in the form *-n* in the modern Aihui dialect, but such a comparison would be anachronistic. Nevertheless, a converb form would be expected because the following word was a verb. Kiyose (1977: 140) assumes that *\*ain* 愛因 may be the same interrogative as Manchu *ainu*, which seems accurate. Kiyose (1977: 144) also mentions a form *adi* 阿的 that has been translated as ‘etc.’ This might correspond to Manchu *udu* ‘how many/much’, but, if true, is closer to northern Tungusic. Notice that Manchu *udu* may also mean ‘several’, which is a bit closer semantically.

It is well-known that Manchu has a very special and aberrant position among Tungusic languages. In fact, the differences are so strong that I have previously put forward the possibility that, in Operstein’s (2015) terminology, Manchu is a contact variety that shows a certain amount of simplification (Hölzl 2012). An additional argument in favor of this hypothesis is the existence of many analyzable interrogatives that consist of either *ai-* or *ya-* together with a noun or, in some cases, another element (Table 5.131). Most forms are not normally written in one word in Manchu, but may be analyzed as compounds. Most dialectal forms that go back to these compounds have not been listed above. Most of these formations are very transparent. In some cases there is an unexpected semantics, such as in *ai-se-me* ‘why’. Hauer (2007) additionally mentions the form *ainam.baha-* ‘to get how’, which is a combination *ainame* ‘how’ and *baha-* ‘to get’. It is highly doubtful that *bi-* in *ai-bi-* is the copula *bi* as claimed by Gorelova (2002:

Table 5.130: Interrogatives in Manchuic (Norman 2013; Zikmundová 2013; Wang Qingfeng 2005; Kim et al. 2008; Zhao Jie 1989); most case forms and some variants are not listed

	Written	Sibe	Aihui	Sanjiazi	Yibuqi
who	we	və	və	wə	və
what	ai	ai	ε	aj	εi
to do what	aina-	ainə-			
which	ya	ya		ja	ja
when	atanggi, ai erinde	aitiŋ, ya erin-t	etiŋ(ŋe), ε(j)irin	ajtiŋ, aj əlɪn	εi t'ɪŋkə
what kind of thing	ai-jaka, ya jaka	ai jaq(ə)	atʂaʂa, εdɛrɛ	aj jakə	εi tʂaʂa
what (is it)	antaka	antq			
which one	ya emken, yaka	yam, yamkə <sup>n</sup>	ja(m)kən, jama		ja əmk'ə
where	aiba-, aibi-, yaba-, ya-, ai-	ε.vi-, yava-, ye-, ai-	εba-, java	jawu- ~ jawə(-)	jaɣa-
how	absi, adaramɛ, ainame	afʂ(e)	əvɕe ~ avɕe	adəlmən	atər(ə)mə
why	ainu, ai turgun	a <sup>n</sup>	an ~ aŋ, ana ~ anə	ai tulxun	ana
how many	udu	ut, yask(ə)	udɛle, ɛdzəxə ~ ɛdzəɣ, ɛdik	udə, udali	utu

219) because it would be impossible to attach a case marker to it, e.g. *ai-bi-de* ‘where’. Most likely it really is a variant of *ba* ‘place’, e.g. *ai-ba-de*. Manchu *ai-ba-* has interesting correspondences in the demonstratives. While the demonstratives have the usual form *e-(re)* ‘this’ and *te-(re)* ‘that’, the correspondence of *ai-ba-* is *u-ba-* ‘here’ and *tu-ba-* ‘there’. A special case is represented by *atanggi*, which seems to be based on *ai* ‘what’. The form may be amalgamated, but, based on the meaning ‘when’, one may suspect a noun meaning ‘time’ to underlie the second part. In fact, *ai-erin-* is such a compound. But there is no word meaning ‘time’ with an adequate form in Manchu and *atanggi* is not synchronically analyzable. Table 5.132 shows all dialectal cognates available.

The Alchuka and Bala forms might indicate a connection with Manchu *antaka*. However, according to Mu Yejun (1988b,a), the *-n-* in one form of Bala and in Alchuka is an innovation. The Kilen form *iətin* has most likely been borrowed but is probably built

5 Survey of the grammars of questions in Northeast Asia

Table 5.131: Analytical interrogatives in Manchu (Hauer 2007; Norman 2013; Hölzl 2015c)

Form	Meaning	Basis	Meaning
ai-ba-, ya-ba-	where	ba	place
ai-erin-	when	erin	time
ya-emu, ya-emken	which one	emu ~ emken	one
ai-hacin, ya-hacin	what kind of	hacin	kind, sort, class
ai-haran	why	haran	reason
ya-ici	whither	ici	direction
ai-jaka, ya-jaka	what sort of	jaka	thing
ai-niyalma, ya-niyalma	what person, who	niyalma	person
ai-se-	to say what	se-	say-
ai-se-me	why	se-me	say-CVB.IPFV > QUOT
ai-turgun	why	turgun	reason
atanggi	when	?	?time

Table 5.132: Cognates of the Jurchenic temporal interrogative; not all variants are shown

Language	Form
Manchu	atanggi
Aihui Manchu	etiŋ(ŋe), etiŋŋi
Sanjiazi Manchu	aitiŋ, aitiŋga, aitiŋŋe
Sibe	aitiŋ
Yibuqi Manchu	ei t'inkə
Bala	a(n)t'əŋniə, at'ani
Alchuka	ant'əŋgi
Kilen	?iətin

on Manchu *ya* instead of *ai*. Given the absence of a fitting word meaning ‘time’ within Jurchenic, it may have been borrowed from another language, not unlike Manchu *hacin* ‘sort’, which according to Benzing (1956: 100), stems from Korean. Middle Korean *enu-cjej* ‘when’ matches the Jurchenic form typologically (‘what-time’), but not formally (§5.7.3). A connection to (Eastern) Old Japanese *tōkyi* 等伎 ‘time’ (Kupchik 2011: 60, 106) seems extremely implausible, but this is the only form I was able to find in surrounding languages that at least looks remotely similar. The interesting proposal by Alonso de la Fuente (2017) that *atanggi* might be related to other words ending in *+nggi* such as *senggi* ‘blood’ is not plausible on semantic grounds and would leave the first part *ata-* unanalyzed. For now, the origin of the interrogative remains obscure.

## 5.11 Turkic

### 5.11.1 Classification of Turkic

The internal diversity of Turkic is much more elaborate than that of, say, Mongolic, but less so than Uralic. Based on Johanson (1998: 81f.; 2006a: 161f.) the languages may roughly be classified as in Figure 5.7. An asterisk indicates that a given language is at least partly spoken in NEA today. Of course, Turkic languages altogether derive from southern Siberia and northern Mongolia (e.g., Yunusbayev et al. 2015), which is why languages such as Turkish and Chuvash will also briefly be addressed. Most languages included here are from the Northeastern or Siberian branch. The Turkic language Yellow Uyghur that is also called Western Yughur (*xībù yùgù yǔ* 西部裕固语 in Chinese) or Sarig Yughur, has to be distinguished from the Mongolic language Shira Yughur or Eastern Yughur (*dōngbù yùgù yǔ* 东部裕固语 in Chinese, §5.8). Yellow Uyghur has no close relation to Uyghur, which belongs to an altogether different branch of Turkic. Despite its name, Fuyu Kyrgyz, spoken in the Heilongjiang province of northeastern China, is more closely related to Yellow Uyghur and the other Abakan Turkic languages than to Kyrgyz as such, which belongs to the Kipchak branch. Similar to the Tungusic language Sibe that was partly relocated to Xinjiang in 1764, Fuyu was brought to Manchuria from the Altai region in the 1750s under emperor Qianlong (see Hu Zhenhua 1986; 1996; Hu Zhenhua & Imart 1987; Janhunen 1996).

*Eynu* (*àinǔ* 艾努 in Chinese) can be considered a truly mixed language. It is “a language that is structurally and grammatically Uyghur, but whose vocabulary is predominantly Persian or Persian-derived.” (Lee-Smith 1996a: 861). Possibly, the origin of *Eynu* lies in its use as a secret language. Because this study is mostly concerned with the grammar of questions, it has been classified as basically Turkic here (see also Tooru et al. 1999 for a discussion). The alternative name *Abdal* has a derogatory meaning and will not be used (see also Ladstätter & Tietze 1994, Wurm 1997). Another special case is *Salar*, which is the result of language contact between Turkic languages from different branches.

*Salar* originated as an Oghuz language and during the course of its speakers’ gradual eastward migration acquired various influences from southeastern- and northwestern-type Turkic languages as well as from non-Turkic ones. It had been assumed earlier that *Salar* is an isolated dialect of Modern Uyghur, mostly on the basis of phonological features such as liquid assimilation and vowel raising. (Hahn 1998: 400)

*Ili Turki* (*tǔ’ěrkè* 土尔克 in Chinese) is a language that is not very well-known and for which only a few descriptions are available. According to Hahn’s (1991: 31) classification it shares properties with both the Kipchak and the Uygur-Karluk branches of Turkic, but it has been tentatively classified with the latter in this study.

5 Survey of the grammars of questions in Northeast Asia

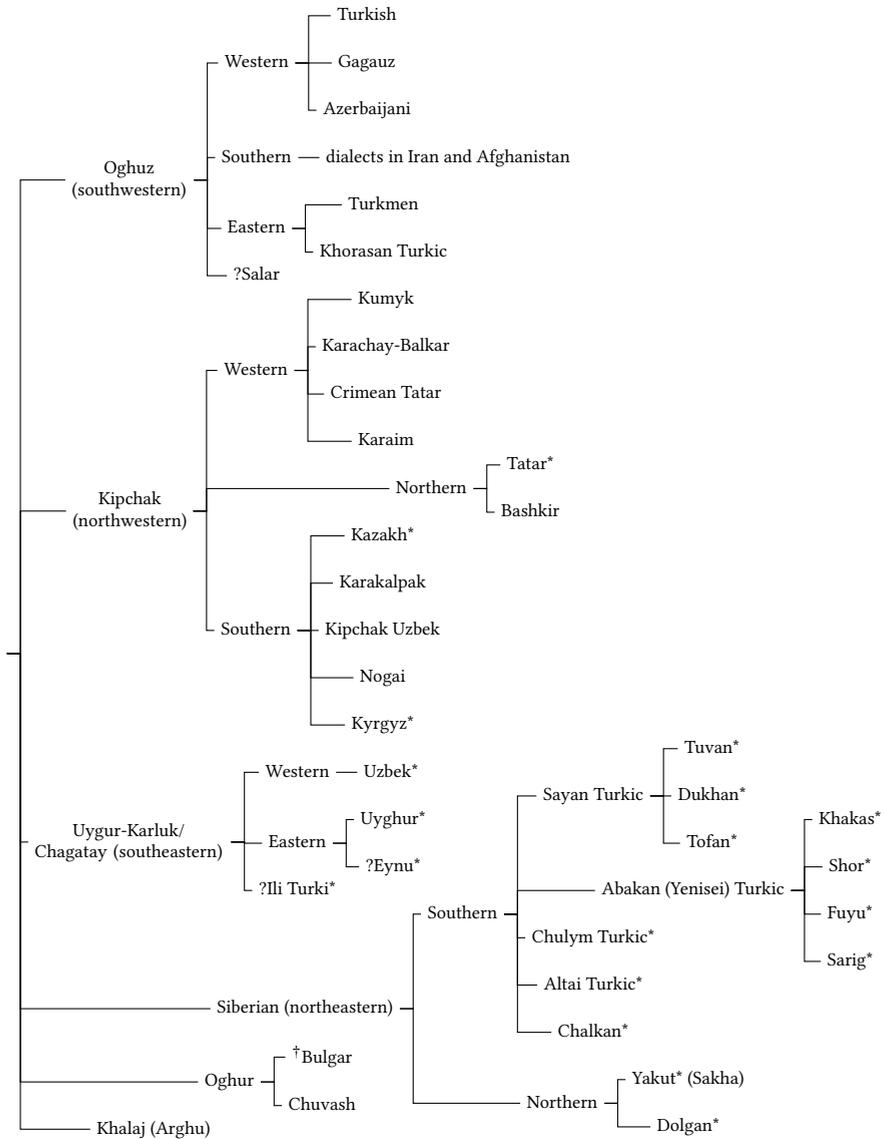


Figure 5.7: Classification of Turkic

### 5.11.2 Question marking in Turkic

The following will briefly describe question marking in the Oghuz language Turkish, which is actually located outside of the Northeast Asian area but may serve as a reference point for the other Turkic languages. Turkish has a sentence-final particle =*mi* that is usually written detached from its host but is best analyzed as enclitic. It has the vowel harmonic variants listed in Table 5.133. There is no variation in the consonant that invariably has the nasal shape *m*.

Table 5.133: Vowel harmonic forms and distribution of the Turkish question marker (Göksel & Kerslake 2005: 22, 251; Landmann 2009: 4, 24)

Preceding vowel	Variant of <i>mi</i>
i, e (front unrounded vowel)	<i>mi</i>
ü, ö (front rounded vowel)	<i>mü</i>
ı, a (back unrounded vowel)	<i>mı</i>
u, o (back rounded vowel)	<i>mu</i>

The morphosyntactic behavior of the Turkish question marker is similar to Proto-Tungusic and some modern Tungusic languages, Middle Mongol or Old Japanese. It is a mobile particle that also marks focus and alternative questions and is also part of several tag question markers. In focus questions it attaches to the element in focus, which receives an additional peak in intonation (Landmann 2009: 24). Content questions remain unmarked. Alternative questions take two markers and have an optional disjunction *yoksa*.

#### (376) Turkish

- a. *Nermin okul-a git-miş=mi?*  
 PN school-DAT go-PFV.EV=Q  
 ‘Has Nermin gone to school?’
- b. *Nermin okul-a=mi git-miş?*  
 PN school-DAT=Q go-PFV.EV  
 ‘Has Nermin gone to school?’
- c. *Cemal okul-a git-ti=mi, (yoksa) git-me-de=mi?*  
 PN school-DAT go-PFV=Q (or) go-NEG-PFV=Q  
 ‘Did Cemal go to school or not?’
- d. *Ev-e ne zaman gid-ecek-sin?*  
 home-DAT what time go-FUT-2SG  
 ‘When will you be going home?’ (Göksel & Kerslake 2005: 257, 254, 259)

Turkish has tag questions marked with the demonstrative *öyle* ‘like this’ or the negative copula *değil* followed by the regular polar question marker *mi* (Göksel & Kerslake 2005: 253).

5 Survey of the grammars of questions in Northeast Asia

(377) Turkish

a. *Cemal bugün okul-a git-me-di, değil mi?*

PN today school-DAT GO-NEG-PFV NEG Q

‘Cemal didn’t go to school today, did(n’t) he?’

b. *Esra Handan-in abla-sı-yımıř, öyle mi?*

PN PN-GEN elder.sister-3SG.POSS-EV.COP so Q

‘So Esra is Handan’s elder sister, right?’ (Göksel & Kerslake 2005: 253)

Similar tag question markers derived from demonstratives and negative copulas are also known from Mongolic (§5.8.2). With this reference point in mind we can now address those Turkic languages spoken in Northeast Asia.

In Salar, the only Oghuz language located in Northeast Asia, polar questions appear to obligatorily take one of several sentence-final question markers. One marker has the vowel harmonic variants *mu*, *mo*, and *mi* and is most likely cognate with Turkish *mi*. Another marker has the two forms *u* and *o* and can be compared with the Mongolian polar question marker  $=\text{(y)UU}$  that has the two vowel harmonic forms  $=\text{(y)uu}$  and  $=\text{(y)oo}$ . It exhibits a short vowel in some, especially Shirongolic languages that are in close vicinity to Salar. The functional difference between the two markers  $=mU$  and  $=U$  remains unclear, but they appear to be mutually exchangeable (Lin Lianyun 1985: 91). The vowel harmonic forms of both particles have a somewhat unclear distribution. For example, all three variants, *mu*, *mo*, and *mi*, are apparently possible in example (378a). The fact that there are different numbers of vowel harmonic variants can be explained by the different origin of the question markers. Both have been reanalyzed as enclitics here.

(378) Salar

a. *sen af if-dži=mu?*

2SG noodle drink-PST.DEF=Q

‘Have you eaten (yet)?’

b. *u ge(l)-mif=u?*

3SG come-PST.INDEF=Q

‘Has (s)he come?’<sup>33</sup> (Lin Lianyun 1985: 90, 91)

Salar content questions are almost always marked with the sentence-final *-i*, which is likely to be related to Tuvan *-il*, Dukhan *-ĭl*, Tofa *-(u)l*, Yakut *-(n)lj*, and Dolgan *-ij*, all of which are restricted to content questions (see below). The forms have all been analyzed as suffixes here, but some might have an enclitic status.

(379) Salar

<sup>33</sup>The meaning of the parentheses in this last example is not entirely clear, but could indicate either optional elements or parts of the verb stem that are lost in combination with the suffixes.

*ana, sen gala va(r)-bur-i?*  
 girl 2SG whither go-FUT-Q

‘Miss, where are you going?’ (Lin Lianyun 1985: 116)

However, in one example Lin Lianyun (1985: 82) has an example with a copula *deri* that possibly contains the question marker, though this was left unanalyzed. The marker *-i* is also absent after verbs with a definite past marking (Lin Lianyun 1985: 71).

(380) Salar  
*sen naŋ-a vulə gel-dzi?*  
 2SG what-DAT for come-PST.DEF

‘Why have you come?’ (Lin Lianyun 1985: 86)

Furthermore, Salar has a question suffix *-du ~ -do* that is directly attached to a verb stem and is said to be connected to the category of evidentiality. It is only used if one has observed oneself that the addressee has finished a certain action.

(381) Salar  
*sen if-du?*  
 2SG drink-Q.SELF.EV

‘(I see that) you have finished drinking?’ (Lin Lianyun 1985: 71)

The very specific meaning as well as the morphosyntactic behavior make it implausible to assume a connection with the Yakut and Dolgan question marker =*duo* ~ =*duu* that we will encounter further below. Like many other languages spoken in China, Salar has borrowed the Mandarin question marker *ba* 吧.

(382) Salar  
*u si-(niyi) gaga-ŋ ira ba?*  
 3SG 2SG-(GEN) e.brother-?2SG.POSS COP Q

‘This is your elder brother, right?’ (Lin Lianyun 1985: 84)

Tatar, Kazakh, and Kyrgyz are the three Kipchak languages spoken in Northeast Asia. I will address them in turn. Tatar as spoken in China has a cognate of the question marker in Turkish and Salar that has the form =*mə* ~ =*mi* (Chen Zongzhen & Yi Liqian 1986: 30). It has likewise been reanalyzed as enclitic here. As opposed to Salar, however, content questions remain unmarked.

(383) Tatar (China)  
 a. *sin qajsə-sə-n ala-səŋ?*  
 2SG which-3SG.POSS-ACC want-2SG  
 ‘Which one do you want?’  
 b. *isməbil abzij tyrkije-ge bar-ma-ban=mə?*  
 PN uncle PN-DAT go-NEG-PTCP.PST=Q

‘Has uncle Ismeril (younger brother of the father) never been to Turkey?’

5 Survey of the grammars of questions in Northeast Asia

- c. *siz ilgeri xaləq bartfasə-va bar-va*n* i-di-giz=**mi**?*  
 2SG.POL before people park-DAT go-PTCP.PST COP-PST-2SG.POL=Q  
 ‘Have you been to the People’s Park before?’ Chen Zongzhen & Yi Liqian  
 1986: 147, 111)

The situation is very similar to Tatar as spoken in Russia, for which there is more information on intonation: In polar questions the question marker is optional and “the pitch is high on the last accented syllable of the clause” (Poppe 1963: 126). Content questions remain unmarked and have either falling intonation with two elements (384b) or first rising and then falling intonation with more elements (384c).

(384) Tatar (Russia)

- a. *zäkiyä kil-de=**me**?*  
 PN come-PST=Q  
 ‘Did Zäkiyä come?’
- b. *närsä bar?*  
 what COP  
 ‘What is it?’
- c. *kičä sez-gä **kem** kil-de?*  
 yesterday 2PL-DAT who come-PST  
 ‘Who came to you yesterday?’ (Poppe 1963: 126)

As compared with the other Turkic languages mentioned in this chapter, questions in **Kazakh** are exceptionally well described (e.g., Geng Shimin & Li Zengxiang 1985; Muhamedowa 2016: 17–24). However, only some aspects of question marking in Kazakh can be included here. For further information, the interested reader is referred to the specialized description of Kazakh interrogative constructions by Zhang Dingjing (1991).

Polar questions are either marked by rising intonation or a particle. Rising intonation has an additional semantic component of surprise, e.g. *ol oqəwǝ?* ‘Is (s)he a student?’ (Zhang Dingjing 1991: 99). The question particle =*MA* is often written detached from its host but probably has the status of an enclitic. But the enclitic is not mobile and thus cannot mark focus questions as in Turkish (Muhamedowa 2016: 17). It marks polar and alternative questions. Alternative questions may take an optional disjunction *ældə*.

(385) Kazakh (China)

- a. *axan da bar-də=**ma**?*  
 e.brother also go-PST=Q  
 ‘Did your elder brother go as well?’ (Geng Shimin & Li Zengxiang 1985: 119)
- b. *ol oqə.wǝ=**ma**, oqə.t.əwǝ=**ma**?*  
 3SG student=Q teacher=Q  
 ‘Is (s)he a student or a teacher?’

- c. *olar faqar-ma-də=ma əlde bar-əη kel-me-di=me?*  
 3PL call-NEG-PST=Q OR.Q go-PST.PFV come-NEG-PST=Q

‘Didn’t they call you or did you (simply) not come?’ (Zhang Dingjing 1991: 99, 104)

According to Muhamedowa (2016: 65) standard Kazakh *əlde* is primarily used in the written language and is restricted to questions. Standard disjunction is expressed with the help of *nemese* ‘or’. This is a distinction also found in Mandarin Chinese *háishì* 还是 (interrogative) versus *huòzhě* 或者 (standard) (§5.9.2.1). In Chinese Kazakh, negative alternative questions either take two markers and an optional disjunction (386a) or, if the second alternative consists of the negative existential exclusively, only the first alternative is marked (386b). In this case, the second alternative simply consists of the negator *dzoq*. This may have been influenced by Uyghur as there are examples with question markers from Kazakhstan (387b). Spoken Kazakh loses its agreement markers if the question marker is present. The sentence *kel-e-di=me?* ‘come-PRS-3SG=Q’ in written Kazakh thus has the spoken equivalent *kel-e=me?* (Muhamedowa 2016: 168). The same phenomenon can be observed in the following alternative question.

(386) Kazakh (China)

- a. *ol kel-e=me, kel-me-j=me?*  
 3SG come-PRS=Q come-NEG-PRS=Q  
 ‘Does (s)he come or not?’
- b. *sarə-maj-də dzaqsə kər-e-siη=be dzoq?*  
 yellow-oil-DAT good see-PRS-2SG=Q NEG  
 ‘Do you like butter?’ (Zhang Dingjing 1991: 104)

(387) Kazakh (Kazakhstan)

- a. *ornija orir-a-siη=ba, orir-ma-y-siη=ba?*  
 place sit-PRS-2SG=Q sit-NEG-PRS-2SG=Q  
 ‘Are you going to sit down at your place or not?’
- b. *ornija orir-a-siη=ba, žoq=pa?*  
 place sit-PRS-2SG=Q NEG=Q  
 ‘Are you going to sit down at your place or not?’ (Muhamedowa 2016: 18)

In sentences with second person singular agreement forms, the particle =MA often has the appearance of a suffix -MI that precedes the agreement suffix. But apparently both constructions are usually possible in such cases.

(388) Kazakh (China)

- a. *nurbek-siη=ba?*  
 PN-2SG=Q  
 ‘Are you Nurbek?’

5 Survey of the grammars of questions in Northeast Asia

b. *nurbek-pi-siŋ?*

PN-Q-2SG

‘Are you Nurbek?’ (Geng Shimin & Li Zengxiang 1985: 120)

The enclitic and the suffix both follow vowel harmony and depend on the preceding consonant, but have different realizations (Tables 5.134, 5.135, see also Kirchner 1998a: 328).

Table 5.134: Realizations of the Kazakh enclitic =MA (Geng Shimin & Li Zengxiang 1985: 119, passim, Zhang Dingjing 1991: passim, and Muhamedowa 2016: 17)

	after back vowels	after front vowels
after <i>r, j, V</i>	=ma	=me
after voiced C	=ba	=be
otherwise	=pa	=pe

Table 5.135: Realizations of the Kazakh suffix -MI preceding second person agreement forms (Geng Shimin & Li Zengxiang 1985: 119, passim; Muhamedowa 2016: 17)

	after back vowels	after front vowels
after <i>r, j, V</i>	-mə	-mi
after voiced C	-bə	-bi
otherwise	-pə	-pi

In spoken but not written Kazakh there is a tendency for the enclitic to lose the vowel harmony in favor of the back vowel variants *ma, ba, pa* (Muhamedowa 2016: 17). Content questions do not have a particle or suffix but exhibit rising intonation (Muhamedowa 2016: 20).

(389) Kazakh (China)

*siz qajda bar-a-siz?*

2SG.POL where go-PRS-2SG.POL

‘Where do you go?’ (Geng Shimin & Li Zengxiang 1985: 68)

A marker specialized to inquire about topics that is restricted to the spoken language takes the form =*ŋl* (i.e. *fə ~ fi*) or =*fe* (Zhang Dingjing 1991: 103, Muhamedowa 2016: 19).

(390) Kazakh (China)

*omar=fe?*

PN=Q

‘What about Omar?’ (Geng Shimin & Li Zengxiang 1985: 121)

A marker with a function similar to tag question markers found in content questions takes the form  $\text{æ}$  and is thus similar to the marker  $a$  in Kalmyk in both form and function (§5.8.2).

- (391) Kazakh (China)  
*munə savan kim ajtə-p ber-di, æ?*  
 this 2SG who tell-CVB give-PST Q  
 ‘Who told you that, eh?’ (Geng Shimin & Li Zengxiang 1985: 131)

In this function  $\text{æ}$  is accompanied by rising intonation and has the specific meaning ‘that the speaker remains baffled by the whole situation of a (certain) circumstance even after giving it much thought.’ (它表示说话人对事物的真象百思不得其解, Zhang Dingjing 1991: 103). Muhamedowa (2016: 19) simply treats standard Kazakh  $\bar{a}$  as a tag question marker and translates it as ‘right?’. It is accompanied with rising intonation.

The question marker in Kyrgyz has the form  $=BI$  (Kirchner 1998b: 346). It takes the form  $=bI$  after  $z, m, n, \eta, l, r, w,$  and  $j$  but the form  $=pI$  after any of the following sounds:  $p, t, k, s, \check{s}, \check{c},$  and  $x$  (Kara 2003: 38). As opposed to Kazakh, there is no variant with an initial nasal. It can attach to variable word classes, which is why it has been analyzed as an enclitic as in Kazakh.

- (392) Kyrgyz (Kyrgyzstan)  
*al kel-e=bi?*  
 3SG come-PRS=Q  
 ‘Is (s)he coming?’ (Kara 2003: 39)

Basically the same situation as in Kyrgyz proper can be observed in Kyrgyz as spoken in China. Here the enclitic has four (as opposed to two in Kazakh) different vowel harmonic variants:  $=Bə, =Bi, =Bu,$  and  $=By$  (Hu Zhenhua 1986: 155).

- (393) Kyrgyz (China)  
*bul kitep=pi?*  
 this book=Q  
 ‘Is this a book?’ (Hu Zhenhua 1979: 89)

In the Teskei dialect of Kyrgyz spoken in Xinjiang (*tiesikai* 铁斯开 in Chinese) the question and agreement markers have the reversed order compared to Kyrgyz proper, but only in the second person.

- (394) Kyrgyz  
*kel-e-sing=bü?*  
 go-PRS-2SG=Q  
 ‘Are you going?’

5 Survey of the grammars of questions in Northeast Asia

- (395) Kyrgyz (Teskei)  
*kel-e-bi-sin?*  
go-PRS-Q-2SG  
'Are you going?' (Makelaike Yumai'erbai 1986: 23)

This unusual morphosyntactic alternation suggests influence from Kazakh or Uyghur. No examples for focus or alternative questions were found in the literature available to me. Content questions are unmarked.

- (396) Kyrgyz (China)  
*qaijda bar-a-səŋ?*  
whither go-PRS-2SG  
'Where are you going?' (Hu Zhenhua 1986: 174)

Apart from the question marker mentioned above, there is a whole range of additional constructions with fine semantic differences. Topic questions take the marker =*tʃI* (i.e. *tʃə ~ =tʃi ~ =tʃu ~ =tʃy*), a cognate of Kazakh =*fI ~ =fe*.

- (397) Kyrgyz (China)  
*men bar-a-mən, siz=tʃi?*  
1SG go-PRS-1SG 2SG=Q  
'I have to go, what about you?' (Hu Zhenhua 1986: 155)

Kyrgyz has a marker *beken* that is a contraction of the polar question marker with a particle *eken*. Similarly, there is a question marker *bejim* that derives from a combination of the question marker with *dejim*.

- (398) Kyrgyz (China)
- a. *al qərɐz beken?*  
3SG PN Q  
'Isn't he Kyrgyz?'
  - b. *al tatar bejim?*  
3SG PN Q  
'(S)he apparently is Tatar, right?' (Hu Zhenhua 1986: 155, 156)

The exact meaning of these forms remains unclear to me.

There are two **Uyghur-Karluk** languages spoken in Northeast Asia as defined here, Uyghur and Uzbek. Uzbek is located for the most part outside of the region but is also spoken by a minority in Xinjiang, which is why it has been included here. In addition, there is the Uyghur-Persian mixed language Eynu that is included in this chapter because of the apparent similarities in question marking to Uyghur. Question marking in **Uyghur** is rather complex. Fortunately we are in possession of descriptions of Uyghur from the 19th century. According to the description by Shaw (1878: 56), Uyghur has a marker =*mu*. Similarly to Kazakh there is a split that in Uyghur depends on the tense affix. The

default position of =*mu* is the very end of the sentence. However, if the non-past marker is present, the question marker has the shorter form *-m* and precedes the agreement marker (Shaw 1878: 56).

- (399) Uyghur  
 a. *qel-ding=mu?*  
 do-PST=Q  
 ‘Did you do?’  
 b. *qel-a-m-san?*  
 do-NPST-Q-2SG  
 ‘Do you do?’ (Shaw 1878: 56)

In fact, Shaw’s explanation is more likely than the one by Tuohuti Litifu (2012: 366) and Abdurehim (2014: 178), who claim that in non-past sentences the marker takes the form *-am* ~ *-äm*. Most likely, the element *-a* ~ *-ä* is a variant of the non-past marker and only *-m* is the question marker.

- (400) Uyghur  
*ätä kel-äl-ä-m-siz?*  
 tomorrow come-ABIL-NPST-Q-2SG.POL  
 ‘Can you come tomorrow?’ (Tuohuti Litifu 2012: 366)

But the descriptions agree that the question marker usually has the invariable form =*mu*.

- (401) Uyghur  
*sän tapşuruq-ni işlä-p bol-duŋ=mu?*  
 2SG homework-ACC make-CVB aux-PST.2SG=Q  
 ‘Have you finished your homework?’ (Tuohuti Litifu 2012: 219)

Shaw (1878: 56) states that, colloquially, =*mu* may have the form =*ma*. But more likely the form =*ma* is a combination of =*mu* and another marker =*a* that expresses astonishment (Tuohuti Litifu 2012: 366).

- (402) Uyghur  
*bayiqi gäp-ni anli-mi-diŋ=ma?*  
 just.now talk-ACC hear-NEG-PST=Q  
 ‘Didn’t you hear what was just said?’ (Tuohuti Litifu 2012: 367)

However, the Lopnor dialect of Uyghur does not show the fusion with the tense marker “when the second person or plural suffixes are attached” and also has a variant *mi*, e.g. *yürü-y-mi-siz* ‘go-NPST-Q-2SG.POL’ (Abdurehim 2014: 178f.). But the author clearly contradicts himself and also gives the following example: *sat-a-m-sän* ‘sell-NPST-Q-2SG (Abdurehim 2014: 208). Polar questions in Uyghur can also be formed by rising intonation alone (Abdurehim 2014: 208). Content questions remain unmarked.

(403) Uyghur

a. *nima bâr?*

what exist

‘What is there?’ (Shaw 1878: 81)

b. *bu nemä?*

this what

‘What is this?’ (Tuohuti Litifu 2012: 367)

Alternative questions have an optional disjunction *yaki* and two obligatory question markers. Most languages with two question markers have the identical question marker on the respective alternatives. The identical part of the two alternatives that is prone to ellipsis usually remains unmarked. Such a situation can also be found in Uyghur (see example 406a below). However, there is a typologically very special situation in Uyghur in which the first of the two markers attaches to the identical element of the two alternatives and has a form different from the second marker.

(404) Uyghur

*bügün kel-a-m-siz (yaki) äti=mu?*

today come-NPST-Q-2SG.POL or tomorrow=Q

‘Do you come today or tomorrow?’ (Tuohuti Litifu 2012: 321)

A similar construction from Kazakhstan has been mentioned by Muhamedowa (2016: 18) but was left unexplained as the Uyghur pattern. In this example, the first question marker retains its shape, but see above on Kazakh for cases in which the agreement marker follows the question marker as in Uyghur.

(405) Kazakh (Kazakhstan)

*samsung al-a-siη=ba, ayfon=ba?*

PN buy-PRS-2SG=Q PN=Q

‘Will you buy a Samsung or an iPhone?’ (Muhamedowa 2016: 18)

The form of negative alternative questions depends on the clause type that determines the type of negator.

(406) Uyghur

a. *sän zadi bar-a-m-sän, bar-ma-m-sän?*

2SG actually go-PRS-Q-2SG go-NEG-Q-2SG

‘Will you actually go or not?’

b. *saηa jan keräk=mu ämä=mu?*

2SG.DAT life need=Q NEG=Q

‘Do you want to live?’

- c. *başqi-lar-din al-yan qärz-ni qaytur-al-a-m-sän yoq?*  
 other-PL-ABL take-P.PFV debt-ACC pay-ABIL-Q-2SG NEG  
 ‘Can you pay the debt you have from the others or not?’ (Tuohuti Litifu 2012: 368)

Notice the absence of the second question marker in the last example with the negative existential *yoq*, a situation already encountered in Chinese Kazakh. There are several additional sentence-final question markers, the meaning of which is not absolutely clear, e.g. *yu, du* (Abdurehim 2014: 208). But *hä* is clearly a question tag, and *ču* must be cognate with the topic question marker from Kyrgyz and Kazakh seen above.

There are only a few descriptions of the mixed language *Eynu*. But the same question marker =*mu* as in Uyghur is present. Consider the following pair of negative alternative questions in Eynu and Uyghur.

- (407) Eynu  
*xani-da mike hes=**mu**, nist=**mu**?*  
 house-LOC goat COP=Q NEG=Q
- (408) Uyghur  
*öj-de öjke bar=**mu** joq(=**mu**)?*  
 house-LOC goat COP=Q NEG(=Q)  
 ‘Are there any goats in the house or not?’ (Wurm 1997: 245)

Lee-Smith (1996a: 860), who has the same example in a slightly different transliteration, does not have the second question marker in Uyghur. As can be seen, the grammatical structure of the Eynu example is nearly identical to Uyghur. However, the lexical stems have a Persian origin. According to Zhao Xiangru & Aximu (2011: 315), there is only one question marker in Uyghur but two in Eynu. Content questions in Eynu are unmarked as in Uyghur.

- (409) Eynu  
*ma jek tʃɛfɛ kɛs kim?*  
 this one eye person who
- (410) Uyghur  
*ma jektʃɛfɛ kifɪ kim?*  
 this one-eyed person who  
 ‘Who is this one-eyed person?’ (Zhao Xiangru & Aximu 1981: 46)

In this example Eynu *kɛs* is of Turkic origin while Uyghur has the same Persian loan-word for ‘one-eyed’ as Eynu. Content questions in Eynu remain unmarked.

- (411) Eynu  
*siz nidʒej-din fɛs\_bol?*  
 2SG.?POL where-ABL come.PST.2SG  
 ‘Where did you come from?’ (Tooru et al. 1999: 31)

5 Survey of the grammars of questions in Northeast Asia

Uyghur has the two tag question markers *šundaq=mu* and *šundaq=qu* that contain the polar question marker or another unidentified marker in combination with a demonstrative.

- (412) Uyghur  
*ular бүгүн кәл-мә-йду,      šundaq=qu?*  
 3PL today come-NEG-3NPST just.so=Q  
 ‘You are not coming today, right?’ (Tuohuti Litifu 2012: 322)

Topic questions such as *sän-ču?* ‘what about you?’ exhibit a marker *-ču* (Tuohuti Litifu 2012: 218) that is probably cognate with the forms seen before in Kazakh and Kyrgyz. The form dubitative *m’ikin* ‘is it?, may it be?’ (Shaw 1878: 81), is cognate with Kyrgyz *beken* and “expresses more of hesitancy between two opinions than the simple *mu*” (Shaw 1878: 56).

One of the least known Turkic languages in NEA is probably **Ili Turki** (Zhao Xiangru & Aximu 1985). There are only a handful of examples for questions. But these are sufficient to illustrate the question marker *=MA*.

- (413) Ili Turki  
 a. *bar-di=ma?*  
 go-PST=Q  
 ‘Did (she) go?’  
 b. *bar-di-q=pa?*  
 go-PST-1PL=Q  
 ‘Did we go?’ (Hahn 1991: 31)

Table 5.136 illustrates that the difference from surrounding Turkic languages is mostly phonological in nature.

Table 5.136: A comparison of three interrogative sentences in six Turkic languages (Zhao Xiangru & Hahn 1989: 278f., slightly modified)

	‘Did he give?’	‘Did you see?’	‘Did we go?’
Ili Turki	berdi=mä?	kördün=bä?	bardıq=pa?
Kazakh	berdi=me?	kördiŋ=be?	bardıq=pa?
Kyrgyz	berdi=mi?	kördün=bü?	bardıq=pi?
Uzbek	berdi=mi?	kordiŋ=mi?	bardıq=mi?
Xinjiang Uzbek	berdi=mi?	kördiŋ=mi?	bardıq=mi?
Uyghur	bärdi=mu?	kördün=mu?	bardıq=mu?

**Uzbek** has a polar question marker *=mi*, which is “accompanied by a rising pitch in the preceding syllable.” (Boeschoten 1998: 373). The same marker is present in Uzbek as spoken in Xijiang (Table 5.136). As in Uyghur, the marker has one form only. It can attach to different word classes.

(414) Uzbek

a. *mümkin=mi?*

possible=Q

‘Is it possible?’<sup>34</sup>b. *futbol oyin-gá bâr-âylik=mi?*

football match-DAT go-VOL=Q

‘Shall we go to the football mtach?’ (Boeschoten 1998: 360, 368)

When second person agreement forms are present, the question marker may either precede or follow the agreement marker.

(415) Uzbek

a. *sây-mi-siz?*

good-Q-2PL

‘Are you well?’

b. *kêlâ-siz=mi?*

come-2PL=Q

‘Will you come?’ (Boeschoten 1998: 373)

By now this phenomenon should be familiar from several languages seen before. As expected, alternative questions take two markers and content questions remain unmarked. Note the alternative question following a content question (§4.4).

(416) Uzbek

a. *mêñâ qaysi.si yaraš-adi, qizili=mi, âqi=mi?*

1SG.DAT which suit-?COP.PST red=Q white=Q

‘Which one suits me, the red one or the white one?’

b. *kêčâ kim-lar kél-di?*

tonight who-PL come-PST

‘What persons have arrived tonight?’ (Boeschoten 1998: 371)

From the **Siberian** branch ten southern and two northern languages are included in this study. Let me first address the southern subbranch. **Tuvan** has the expected polar question marker *be*. But there is a marker *-il* that appears to be developing into a content question marker, e.g. *kim-il?* ‘Who is it?’ (Anderson & Harrison 1999: 88f.). The marker is optional as there are also unmarked content questions.

(417) Tuvan

a. *xlep bar=be?*

bread COP=Q

‘Is there bread?’

---

<sup>34</sup>This may also be pronounced as *mümkimmi*.

5 Survey of the grammars of questions in Northeast Asia

b. *kayiin kel-gen siler?*

whence come-PST.I 2PL

'Where have you come from?' (Anderson & Harrison 1999: 69, 28)

The situation in **Dzungar Tuvan** spoken in China is almost identical. But the polar question marker is =BA with vowel harmony as well as variation of the consonant, both of which are absent in Tuvan proper.

(418) Tuvan (Dzungar)

a. *ol gel-di=be?*

3SG come-PST.II=Q

'Has (s)he come?'

b. *sen cazan dzor-ur-sen?*

2SG when leave-?FUT-2SG

'When will you leave?'

c. *suu-bəs-təŋ bir.si qajda-l?*

water-1PL.POSS-?ABL some where-Q

'Where has some of our water gone?' (Wu Hongwei 1999: 113, 144, 54)

Wu Hongwei (1999: 146) mentions that questions may exhibit a certain question intonation, but leaves open the details.

Basically the same pattern seen in the two varieties of Tuvan above can also be found in **Dukhan**. Again, the polar question marker =BA has more varieties than in standard Tuvan and the content question marker is optional.

(419) Dukhan

a. *batə gel-gen-be?*

PN come-POST-Q

'Has Bat arrived?'

b. *gim gel-gen?*

who come-POST

'Who has come?'

c. *am pis ganjaar-əl?*

now 1PL which.V.INTRA.LF-Q

'And now what should we do?' (Ragagnin 2011: 193, 131, 188)

In Dukhan, the content question marker -(*l*) is said to have an additional intensifying function, which would explain its absence in some sentences in Tuvan as well.

A negative tag question in Dukhan has the form of a negative copula followed by the polar question marker.

- (420) Dukhan  
*gel-gen emes=pe?*  
 come-POST NEG=Q  
 ‘(S)he arrived, didn’t (s)he?’ (Ragagnin 2011: 187)

This pattern has the form *eves=be* in Tuvan (Harrison 2005: 23) and *emes=pe* in Dzungar Tuvan (Mawkanuli 2005: 209). In Sarikoli there is a similar question tag *emes he* with a question marker that is said to also occur in Uyghur (Tooru et al. 1999: 31-32). There is an areal connection of this construction to similar constructions in Mongolic (e.g., Mongolian *bish=uu*, §5.8.2), as well as other Turkic languages (e.g., Turkish *değil=mi*, see above). Tag questions in Tuvan have the final element *ale*, Dukhan has *hala ~ harən* (Ragagnin 2011: 187).

**Tofa** (previously also called Karagas), like Tuvan, has the invariable marker *=be* (Schönig 1998: 414). According to Castrén (1857b: 71);, there is a variation between two forms *-bè ~ -pè* that can attach to different word classes, which is why they can be considered enclitics. Alternative questions, for which no example was given, take the same marker on each alternative. Content questions have a marker *-(u)l* that is cognate with Dukhan *-(ĭ)l* etc.

- (421) Tofa
- a. *onu soodap beer=be?*  
 3SG.ACC say.CVB OBJ.VERS.P/F=Q  
 ‘Should I say it again (for you)?’ (Anderson 2001: 260)
  - b. *ad-iŋ qum-ul?*  
 name-2SG.POSS who-Q  
 ‘What is your name?’ (Schönig 1993: 199)

Question marking in Tofa is thus almost identical to Tuvan.

For **Khakas** (previously also called Koibal), Castrén (1857b: 71) mentions a polar question marker *-BA* (i.e., *-ba*, *-bä*, *-pa*, *-pä*) that should be considered an enclitic as well. There are additional variants with an initial nasal not mentioned by Castrén (Anderson 1998: passim). Alternative questions take two markers and content questions are unmarked.

- (422) Khakas
- a. *sırer xakas-ta-p čooxta-pča-zar=ba?*  
 2PL PN-V-CVB speak-PRS-2PL=Q  
 ‘Do you (pl.) speak Khakas?’
  - b. *kem pıl-er ani?*  
 who know-FUT 3SG.ACC  
 ‘Who knows him/her?’ (Anderson 1998: 87)

The following polar and focus questions were given to me in January 2016 by a native Khakas living in Germany with the help of several Khakas speakers in Russia. The transcription and analysis roughly follow Anderson (1998).

(423) Khakas

- a. *sin taŋ.da, škola par-ča-zŋŋ?*  
 2SG tomorrow school go-PRS-2SG  
 ‘Are you going to school tomorrow?’
- b. *sin taŋ.da, škola par-ča-zŋŋ?*  
 2SG tomorrow school go-PRS-2SG  
 ‘Is it you who is going to school tomorrow?’
- c. *taŋ.da, sin škola par-ča-zŋŋ?*  
 tomorrow 2SG school go-PRS-2SG  
 ‘Is it tomorrow that you are going to school?’
- d. *sin škola, taŋ.da par-ča-zŋŋ?*  
 2SG school tomorrow go-PRS-2SG  
 ‘Is it to school that you are going tomorrow?’

No polar question marker was present. Detailed information on intonation is not available to me, but focus is also clearly marked with word order, especially sentence initial position. In one instance the focused element stands in second position.

However, Khakas has a special interrogative verb ending *-žAŋ* found in content questions that expresses “semi-rhetorical utterances like ‘how is it possible that...?’, when the speaker believes what they are questioning to in fact not be possible or appropriate” (Anderson 1998: 38). It often combines with interrogatives such as *noya* ‘why’, *xaydi* ‘how’ or *xaydar* ‘whither’.

(424) Khakas

- xıyır-yan čir-zer xaydi par-ba-žAŋ?*  
 invite-PPST place-ALL how go-NEG-Q  
 ‘How is it possible not to go where one was invited?’ (Anderson 1998: 38)

Given that *-žAŋ* originally expressed the habitual past, a connection to Samoyedic languages seems possible (§5.12.2). In Enets and Nenets, for example, the suffix *-sa* used to be a past tense marker but acquired an interrogative meaning as well. A difference is that it can be found in both polar and content questions.

According to Hu Zhenhua & Imart (1987: 29) Fuyu has a question marker =BA that has at least eight different realizations, *ba*, *bī*, *pa*, *pī*, *ma*, *mī*, *βa*, and *βī*. They write the question marker attached to the preceding word, but it has been analyzed as enclitic here as it can be attached to a verbal or non-verbal host. Content questions remain unmarked, e.g. *ol gīm?* ‘Who is (s)he?’ (Hu Zhenhua & Imart 1987: 37).

(425) Fuyu

- ol sin-iŋ balaŋ=ma?*  
 3SG 2SG-GEN son=Q  
 ‘Is he your son?’ (Hu Zhenhua & Imart 1987: 29, 37)

In Sarig Yughur the polar question marker has the form =*mi* and an older variant =*pi* after plosives as well as a stressed version =*mʊ* (Roos 2000: 152). It has the reduced form -*m* after the past tense suffix -(<sup>h</sup>)*tī* and the evidential -*tī*.

- (426) Sarig Yughur
- a. *yahs=mu?*  
good=Q  
'How are you?' (a greeting)
- b. *sen uzi<sup>h</sup>tī-m?*  
2SG sleep-PST-Q  
'Did you sleep?' (Roos 2000: 152)

The variant =*mu* as well as the suffix -*m* indicate strong influence from Uyghur. The difference from Uyghur only concerns the lack of a person marker following -*m*, but this category is altogether absent from Sarig Yughur (Roos 2000: 100). In the first two examples =*mi* is also possible. According to Chen Zongzhen (1982: 76) the question marker has the slightly different form =*be*, =*me* ~ -*m*, but no example of the former has been given. The distribution of =*me* and -*m* confirms Roos' description. (Negative) alternative questions take two question makers.

- (427) Sarig Yughur
- kʊ ki-ʃtī-m ki-yīmstī-m?*  
3SG come-FUT.EV-Q come-FUT.NEG.EV-Q  
'Will (s)he come or not?' (Roos 2000: 152)

Roos mentions another example of an alternative question with a very unusual structure. If there is no mistake, the sentence has a question marker at the end of the sentence *in addition* to the two markers already present. Moreover, there is a disjunction that *follows* both alternatives.

- (428) Sarig Yughur
- sen puyin=mi, tayin=mi ta<sup>h</sup>qī c<sup>h</sup>unʃaŋ-qa kuŋcuola-ʃ=mi?*  
2SG today=Q tomorrow=Q or chief-DAT work-FUT=Q  
'Will you work for the chief today or tomorrow?' (Roos 2000: 152)

In Shor the question marker has the variants =*ba*, =*be*, =*pa*, =*pe*, =*ma*, =*me* (Donidze 1997: 505), which is identical to Altai Turkic. Polar questions take one (Nevskaja 2000: 291), content questions are unmarked, alternative questions take two markers.

- (429) Shor
- a. *kem kel-du?*  
who come-PST  
'Who came?'

5 Survey of the grammars of questions in Northeast Asia

- b. *ak kazyn özer=be čok=pa?*  
white ?birch ?grow=Q NEG=Q

‘Is the white birch tree growing or not? (Donidze 1997: 505)

The question marker in **Altai Turkic** has the same variants =*ba*, =*be*, =*pa*, =*pe*, =*ma*, =*me* as in Shor (Baskakov 1997: 183). In addition there is an enclitic =*na* ~ =*ne*. The difference in meaning has not been specified, but the latter marker may have dubitative meaning.

(430) Altai Turkic

- a. *albyŋ=ba?*  
?receive=q

‘Have you received?’

- b. *olor kel-gen=ne?*  
3PL come-?P.PST=Q

‘(I wonder whether) they came?’ (Baskakov 1997: 183)

Content questions appear to be unmarked, e.g. *kajda sin?* ‘Where are you?’ (Baskakov 1958a: 104).

The newly identified South Siberian Turkic language **Chalkan** offers a picture that strongly resembles its closely related languages such as Altai Turkic. Polar questions take a sentence-final question marker and content questions remain unmarked.

(431) Chalkan

- a. *eme t’etire uyuqta-p-sin=mi?*  
now till sleep-PRS-2SG=Q

‘Have you been sleeping till now?’

- b. *t’e, qančïn-dan kečirsin?*  
well how.much-ABL ferry.PRS.2SG

‘Well, how much is it for you to ferry us (over the river)?’ (Nevskaja 2014: 78)

- c. *kem-niŋ tiž-i apaaš’ aa-ŋ=ma meeŋ=me?*  
who-GEN tooth-3SG.POSS rather.white 3SG-GEN=Q 1SG.GEN=Q

‘Whose teeth are especially white, his/hers or mine?’ (Erdal et al. 2013: 98)

The exact variants of the question marker remain unclear.

Little material is also available for questions in **Chulym Turkic** (also called Ös), but some material based on fieldwork has been published by Harrison & Anderson (2003) and Anderson & Harrison (2006). Unfortunately, their data do not contain an example of a polar question, but there are several content questions as well as one open alternative question in which the second part is *ili qajduy?* ‘or what?’ (Anderson & Harrison 2006: 57). The disjunction *ili* as well as the accompanying construction has been borrowed from Russian. Content questions are unmarked.

- (432) Chulym  
*kajnaar bar-eydi-ŋ?*  
 whither go-PRS-2SG  
 ‘Where are you going?’ (Harrison & Anderson 2003: 250)

There is no information on intonation either. Anderson & Harrison (2004: 184) mention an example of a polar question provided by a Middle Chulym speaker, i.e. *uluy=be?* ‘Was it big?’. But apparently, the sentence represents an example of code mixing with Tuvan. The same question marker also exists in Chulym, but has the form *=ba* according to Birjukovich (1997).

- (433) Chulym  
*män par-ad-ym=ba?*  
 1SG go-PRS-1SG=Q  
 ‘Do I go?’ (Birjukovich 1997: 496)

Perhaps the form *=be* was simply borrowed from Tuvan. However, according to the data by Li Yong-Söng et al. (2008: 97) in Middle Chulym, apart from *=ba*, there is a second variant *=bä* that follows front vowels and could be identical to *=be* seen above.

- (434) Chulym (Middle)  
*seen eer-iŋ äp-te=bä?*  
 2SG.GEN husband-2SG.POSS house-LOC=Q  
 ‘Is your husband at home?’ (Li Yong-Söng et al. 2008: 98)

But interestingly, the question marker does not necessarily stand sentence-final, but may be followed by other elements. The description by Li Yong-Söng et al. (2008) is insufficient for a clear analysis. The example is not a focus question in which the mobile question marker attaches to the focused element. Instead, one element from the sentence simply follows the predicate, which is the host for the question marker. Perhaps this is a focus question in which focus is expressed with the help of word order. The following sentence was translated as ‘Do you have good news?’, which seems to be inadequate.

- (435) Chulym (Middle)  
*siler-niŋ čaqši=ba vest’?*  
 2PL-GEN good=Q news  
 ‘Is your news good?’ (Li Yong-Söng et al. 2008: 210)

In Yakut and Dolgan, the two northern Siberian Turkic languages, there is an enclitic *=duo ~ =duu* (Stachowski 1993: 83; Ebata 2011: 197) that marks polar and alternative questions and another marker *-(n)ij* (Ebata 2011: 197) or *-ly* (Stachowski & Menz 1998: 423), showing vowel harmony, that is found in content questions. The first marker is unique among Turkic languages. It can also be found in Kolyma Yukaghir (Nagasaki 2011: 245) but does not exist in Tundra Yukaghir (see §5.14.2).

5 Survey of the grammars of questions in Northeast Asia

(436) Yakut

- a. *kyærəbej uruua-tuu-n ist-e-bin=duo?*  
 lark song-3SG.POSS-ACC hear-PRS-2SG=Q  
 ‘Do you hear the song of larks?’
- b. *obo-ŋ kuuus=duu uol=duu?*  
 child-2SG.POSS daughter.COP.3SG=Q son.COP.3SG=Q  
 ‘Is your child a daughter or a son?’
- c. *xanna baar-gum-uj?*  
 where exist-COP.2SG-Q  
 ‘Where are you?’ (Ebata 2011: 197)

(437) Dolgan

- a. *taahik-ka uu baar=duo?*  
 washbasin-DAT water EX=Q  
 ‘Is there water in the washbasin?’
- b. *tugunan bar-ia-ŋ-iy?*  
 what.INST go-FUT-2SG-Q  
 ‘What will you go with?’ (Li Yong-Söng 2011: 62, 90)
- c. *ikki=duo üs=duo čahy?*  
 two=Q four=Q hour  
 ‘for two or four hours?’ (Stachowski 1993: 83)

Table 5.137 Summarizes question marking in Turkic languages in polar and content questions. Except for Salar, Tuvan, Dukhan, Yakut, and Dolgan, the Turkic languages in the sample have unmarked content questions. Apart from phonological differences, there is very little variation in polar question markers. Only Yakut and Dolgan deviate from the rest of the languages, possibly due to Yukaghiric influence. The content question marker in these Turkic languages has an areal connection to the so-called Mongolic “corrogrative” particle \**büi* (§5.8.2). Similar to Mongolic, the Turkic markers may have their origin in a copula form. For Tuvan, Anderson & Harrison (1999: 88) speculate that *-Il* is a quasi-copula that derived from the demonstrative *ol* ‘that’, which is also the source of the third person singular pronoun (see also Ragagnin 2011: 188). In fact, a development from demonstrative to copula is widely attested, for example for Chinese *shì* 是 or Russian *eto/это*. No Turkic language from the sample marks polar and content questions in the same way. Those languages for which sufficient data are available indicate that alternative questions are usually marked in the same way as polar questions but often exhibit an optional disjunction as well. Altogether there is little variation in polar question markers among Turkic languages. Except for Yakut and Dolgan *=duo* ~ *=duu* and maybe Salar *=U*, all polar question markers listed in Table 5.137 appear to be cognate with Turkish *=mI*. The major difference lies in how many variants the question marker has in a given language. As has often been observed, a very likely origin of the question

marker lies in negation, e.g. Turkish *-mA* or Old Turkic *-mA* (Erdal 1998: 151). The polar question marker was already present in Old Turkic and in Chagatay as *=mU*, with content questions unmarked (Erdal 1998: 152; Boeschoten & Vandamme 1998: 171, 175).

Table 5.137: Summary of question marking in Turkic

	PQ	CQ
Altai	=MA#	-
Chagatay	=mU	-
Chalkan	=MI#	-
Chulym (Middle)	=bA#	-
Dolgan	= <b>duo</b> #	V-ij
Dukhan	=BA#	V-(Ī)l
Eynu	=mu#	-
Fuyu	=BA#	-
Ili Turki	=MA#	?-
Kazakh	=MA#, V-MI	-
Khakas	=BA#	-
Kyrgyz	=BI#	-
Old Turkic	=mU	-
Salar	=mU#, =U#	V-i
Sarig Yughur	=Mi#, =mu# ~ V-m	-
Shor	=MA#	-
Tatar (Chinese)	=mI#	-
Tofa	=be#	V-(u)l
Tuvan	=be#	V-Il
Tuvan (Dzungar)	=BA#	V-l
Uyghur	=mu# ~ V-m	-
Uzbek	=mi#	-
Yakut	= <b>duo</b> # ~= <b>duu</b> #	V-(n)ij

However, in **Chuvash**, the only extant **Oghur** language and the most aberrant Turkic language, there are three question markers that are usually written attached to the preceding word, which takes an intonational peak.

## (438) Chuvash

- a. *vāl χula-na kay-nă-i?*  
3SG city-DAT go-PST-Q  
'Did (s)he go to the city?'
- b. *vāl χula-na kay-nă-im?*  
3SG city-DAT go-PST-Q  
'Did (s)he really go to the city?'

- c. *vāl χula-na kay-nă-ši?*  
 3SG city-DAT go-PST-Q  
 ‘You mean (s)he went to the city?’ (Clark 1998: 450)

While *-i* marks plain polar questions, *-im* is said to express uncertainty or surprise, and sentences marked with *-ši* express doubt and are in need for further confirmation (Clark 1998: 450). Alternative questions take two plain polar question markers and an optional disjunction *e*. Content questions are usually unmarked but may also take the marker *-ši*. In example (439a) an alternative question follows a content question (§4.4).

(439) Chuvash

- a. *esir āšta purăn-a-tăr, xula-ra-i (e) yal-ta-i?*  
 2SG.POL where live-PRS-2SG.POL city-LOC-Q or village-LOC-Q  
 ‘Where do you live, in the city or in the village?’
- b. *muzey āšta-ši?*  
 museum where-Q  
 ‘Where is the museum?’
- c. *esir āšta śural-nă?*  
 2SG.POL where be.born-PF  
 ‘Where have you been born?’ (Landmann 2014: 34, 62)

These data show a pattern markedly different from those Turkic languages located in NEA. While double marking and disjunctions are attested in other Turkic languages as well, there is no formal match to Chuvash. Furthermore, no other Turkic language investigated here has a question marker that can appear in both polar and content questions. However, there was an additional dubitative marker (Old Turkic *erki*, Chagatay *é(r)ken* ~ *é(r)kin*) that could also appear in content questions.

Descriptions of Turkic languages rarely explicitly state the syntactic behavior of interrogatives. Often, for instance in Uzbek and Kazakh, interrogatives are found in focus position in front of the verb (Boeschoten 1998: 373; Muhamedowa 2016: 20).

### 5.11.3 Interrogatives in Turkic

Turkic languages have both KIN- and K-interrogatives. The Proto-Turkic interrogative meaning ‘who’ has been reconstructed as *\*kem* by Róna-Tas (1998: 74) or as *\*kām* for Oghur (Chuvash *kam*), and *\*kim* otherwise by Schönig (1999: 64, 69). Oghuz languages, but not Salar, differ from other Turkic languages in having derivations from the interrogative *ne* ‘what’ instead of *\*qay-* for locative forms (Schönig 1999: 66). In fact, in modern Turkish most interrogatives start with an *n-* (Göksel & Kerslake 2005: 251). In Turkic, however, the interrogative meaning ‘what’ (Turkish *ne*) was originally the only native word starting with an *n-* (e.g., Róna-Tas 1998: 74). This phonotactic anomaly might indicate borrowing from another language. Comparable forms in Northeast Asia exist in Ainuic, Eskaleut, Japonic, Sinitic, and Yukaghiric languages, but none is a very likely source of the Turkic interrogative. Stachowski’s (2015) claim that *\*ne* derives from a

Uralic demonstrative is not very convincing and deserves further evidence. The Uralic language Selkup has several interrogatives that have a Turkic appearance and were probably borrowed (§5.12.3).

Many Turkic languages have a difference between a plain velar plosive in ‘who’ and an uvular plosive in other interrogatives. A similar phenomenon is known from some Mongolic languages (§5.8.3). However, the distribution of the resonances *k*~, *n*~, and *q*~ more closely resembles Yukaghiric languages (§5.14.3). Table 5.138 gives an overview of cognates of five Turkic interrogatives. In most languages the interrogatives ‘which’ and ‘where’ are analyzable synchronically and thus do not qualify as so-called *basic-level* interrogatives. Siberian Turkic languages contain an innovative form combining the meanings ‘how’ and ‘which’ that exhibits variation between *-n-* (e.g., Tuvan *kan-dig*) and *-y-* (e.g., Yakut *χay-daχ*). This variation is already attested in Old Turkic *kañu* ~ *kayu*. There is some overlap with another group of languages that exhibit an older derivation with a suffix *-si* that is even attested in Chuvash *xă-š(ě)* and Khalaj *qāni(-si)*. Apart from the latter, these are usually based on the variant with *-y-* (e.g., Tuvan *kay(i)-zi*).

Stachowski (1990; 2015; p.c. 2016) has quite convincingly argued that Yakut *tuox* ‘what’, apart from Dolgan *tuok* ~ *tuogu*, seems to have no cognates in other Turkic languages. However, while there are certain phonological problems, there might actually be cognates in other Siberian Turkic languages (Table 5.139). Note, first of all, that there are additional forms that must be connected with *tuox* in Yakut. These are the forms meaning ‘why’ (*toyo*) and ‘how many’ (*töhö*). Despite their apparent differences, they share a resonance in *t*~, which suggests a common origin. Second, in both Yakut and Dolgan, as well as the other Siberian Turkic languages, these forms are by and large restricted to these three functions. Third, apart from the question of whether there is a sound law that connects the Yakut and Dolgan forms with the interrogatives from the other languages, they certainly have a similar overall form. Fourth, the fact that all languages are from the same branch of Turkic makes it very plausible to seek a common origin for this anomaly. Fifth, the forms meaning ‘what’ and ‘why’ in all languages have a different vowel quality than the form meaning ‘how many’.

Tuvan *čü-den* and Karagas (Tofa) *tü-dän* ‘why’ contain an ablative instead of a dative (Anderson & Harrison 1999: 28; Castrén 1857b: 163). The derivation of some forms such as Chalkan *t’üg(g)erek*, *t’ugerek*, *t’uriq* ‘what’ remain unclear to me. There is also an interrogative verb ‘to do what’ in Tofa (*čoon-*) and Chalkan (*t’uвет-*).

Stachowski (2015) tried to connect the Yakut form with an Uralic demonstrative stem, which is possible but unlikely from a typological perspective. Interrogatives and demonstratives may share paradigmatic similarities and may also grammaticalize into similar categories such as relatives, but demonstratives do not usually develop into interrogatives. That this change occurred during borrowing, which in itself is not the most likely scenario, is not very plausible either. The only similar form in terms of both meaning and form that I was able to find in NEA can be found in Iranian languages (e.g., Sogdian (*ə*)*ču* ‘what’). However, a connection in terms of borrowing seems too far-fetched. According to Stachowski (2015: 85), *tuox* goes back to *\*to-ok*, in which the suffix is an intensifier. Possibly, the suffix can be compared with Altai *d’u-γ* ~ *ču-γ* and Chalkan *t’ü-γ* ~ *t’u-γ/g*. The other Turkic languages with the unusual interrogative show a palatalized consonant

5 Survey of the grammars of questions in Northeast Asia

Table 5.138: Cognates of five Turkic interrogatives; Chaghatay taken from Boeschoten & Vandamme (1998: 171, 173), Chuvash from Landmann (2014: 32f.), Dukhan from Ragagnin (2011: 94), Khalaj from Doerfer (1988: 107f.), and South Siberian Turkic partly from Schönig (1998: 410); see the rest of this chapter for additional variants

	who	what	which	when	where
Turkish	kim	ne	hangi	ne zaman	nere, hani
Chuvash	kam	mě̃n	xã-š(ě̃)	xãšan	ãšta
Khalaj	kim, kũm	nã̃y	qãni(-si)	qa' çãn	nĩ <sup>c</sup> rã
Chaghatay	kim	ne	qay(u)(-si)	qaçan	qanĩ, qan-da, qay-da
Old Turkic	kãm, kim	nã	kañu, kayu	kaçan	kañu-da, kayu-da
Salar	kam	nang	ga-si	gaqiang	ga-da
Tatar	kim	ni	qaj(-sə)	qajtʃan	qaj-da
Kazakh	kim	ne, nemene	qaj-sə	qafan	qaj-da
Kyrgyz	kim	emne	qaj-sə	qatʃan	qaj-da
Uyghur	kim	nemä	qay-si	qaçan	nä
Ili Turki	kim	nemä	qay-sı	qaçan	
Kashgar Uyghur	kim ~ tʃim	nime	qa(j)-si	qatʃan	nεε
Eynu	kim	nime	qaj-si	qatʃan	nε
Uzbek	kim	nimä	qay-si	qaçan	qani, qayér-dä, qayâ-tä
Tuvan	kim	čüü	kay(i)(-zi), kan-dig	kažan	kay-da
Dzungar Tuvan	cəm	dzy-dzi-mε	caj(ə)-sə, can-dəy	cažan	caj-da
Tofa	qum	čü	qaj-sy, qan-dyy		qaj-da
Karagas (Tofa)	kəm, kum	tü	kan-deg	kašan, kähän	kai-da
Dukhan	gim	jü(ü)	gae		
Khakas	kem	nime	xay(-zi), xay-day	xayžan	xay-da
Koibal (Khakas)	kem, kim	nô, nêmä, nime	kai-ze, kai-dak	kađen	kai-da
Fuyu	gim	n <sup>y</sup> em	cay-zı, ca-dah	cajan	cay-da
Shor	kem	noo	kaj(y)	qaçan	kaj-da
Sarig Yughur	k <sup>h</sup> im	ni	qay-si	qa <sup>h</sup> tan	qay-ta
Altai	kem	d' u-γ ~ ču-γ, ne	kan-dyj	kaçan	kaj-da
Chalkan	kem	t' u(u), t' ü-γ ~ t' u-γ/g, ne	qan-diy, qan-du(γ)	qažan	qay-da
Chulym	kim	tʃ <sup>o</sup> , nöömä	qay-diy	qaçan ~ qajan	qay-da
Yakut	kim	tuo-χ	χaya, χay-daχ	χahan	χanna
Dolgan	ki(i)m	tuo-k ~ tuo-gu	kaja, kaj-da(a)k	kahan ~ kagan	kanna

Table 5.139: A tentative list of cognates of a possible interrogative stem in Siberian Turkic (except for Abakan); not all forms are shown

	what	why	how many
Altai	d'u-γ ~ ču-γ		
Dukhan	ǰü(ü)		
Chalkan	t'u(u), t'ü-γ ~ t'u-γ/g		
Chulym	tʃ'o		
Karagas (Tofa)	tü	tü-gä, tü-dän	teše, tehe
Tofa	čü		čehe
Tuvan	čüü	čü-ge, čü-den	čeže
Tuvan, Dzungar	dzy-dzime	dzy-ge, dzy-nen	dzeze
Dolgan	tuo-k ~ tuo-gu	to-go ~ tuo-go	töhö
Yakut	tuo-χ	to-γo	töhö

instead of *t~*. Most likely this is the result of the following high vowels. The reason for the apparent irregular development in Yakut *tuoχ* and Dolgan *tuok* is not perfectly clear, but one possibility would be an analogy to the negative existentials, Yakut *suoχ* and Dolgan *huok*, that have a regular development. Possibly, the question marker =*duu* ~ =*duo* in Yakut and Dolgan derives from the same source (INT > Q), but this likewise remains somewhat speculative. More research by experts of these languages will be necessary to clarify these points.

For Old Turkic, several partial interrogative paradigms are attested. Table 5.140. compares some of them with Sarig Yughur, for which paradigms were given by Roos (2000). Apart from phonological changes there are only minor differences between the two languages, which illustrates the relatively young age of Turkic.

For reasons of space, paradigms will not be given in detail for other languages throughout this section.

Table 5.140: Old Turkic interrogative paradigms (Erdal 2004: 211) in comparison with Sarig Yughur (Roos 2000: 87)

	who		which	
	Old Turkic	Sarig Yughur	Old Turkic	Sarig Yughur
NOM	<i>kim, käm</i>	<i>k<sup>h</sup>im</i>	<i>kayu, kañu</i>	<i>qay-si</i>
ACC	<i>kim-ni</i>	<i>k<sup>h</sup>im-ni</i>	<i>kayu-ni</i>	<i>qay-si-n</i>
GEN	<i>kim-(n)iñ</i>	<i>k<sup>h</sup>im-niñ</i>	<i>kayu-nuñ</i>	-
DAT	<i>kim-kä, käm-kä</i>	<i>k<sup>h</sup>im-ki</i>	<i>kayu-ka</i>	<i>qa-ya</i>
ABL	-	<i>k<sup>h</sup>im-tin</i>	<i>kayu-din</i>	<i>qay-tan</i>
LOC	<i>kim-tädä</i>	-	<i>kayu-da, kañu-da</i>	<i>qay-ta ~ qan-ta</i>

5 Survey of the grammars of questions in Northeast Asia

Let us now consider interrogatives from individual modern languages. The order will be roughly the same as in §5.11.2, starting with the only Oghuz language Salar (Table 5.141). If available, several descriptions are contrasted for any given language. In some cases only a selection of forms is given.

Table 5.141: Salar interrogatives (Ma Quanlin et al. 1993: passim; Lin Lianyun 1985: 52, 109, 136, passim); some questionable variants were excluded

	Ma Quanlin et al.	Lin Lianyun
who	kam	kem
where	gada	cada
whither		cala
whence		caden
when	gaqiang, gahao	cadzan, cahal
which	gasi	cajsi
for what	naima	
what, how etc.	naiqi	
how	naiqiu	
what	nang	naŋ
why	neigei	neye
how (long)	neisi	
how much/many	neisiqu	nehdze

Similarly to Turkish and Tatar, relatively many of the interrogatives start with *n*-. Only *kam* ~ *kem* ‘who’ has an initial *k*, while all other forms start with *g* ~ *c*. Lin Lianyun (1985: 76) in addition mentions an interrogative verb *naxcur* ‘to do what’ that is claimed to be a contraction of *naŋ et-gur* with the definite future marker. With other verb endings the periphrastic construction is still present.

- (440) Salar  
*sen naŋ et-bər-i?*  
 2SG what do-PROGR.INDEF-Q  
 ‘What are you doing?’ (Lin Lianyun 1985: 86)

Interrogatives from the Kipchak languages Tatar, Kazakh, and Kyrgyz are listed in Table 5.142. The languages all have a similar resonance pattern with the interrogative ‘who’ being the only one that does not show *q*- or *n*-. Tatar is probably no exception, although in Cyrillic transcription the forms all start with *κ*-. This has been transliterated with *q*- before an *a*. Only Kyrgyz *emne* is an exception from the resonances. Most likely it is an *allegro* form that developed from a form similar to Kazakh *nemene*. For comparative purposes, Table 5.142 also contains forms from Tatar proper, transliterated from Cyrillic.<sup>35</sup> Tatar *nerse* derives from *ni erse* (Chen Zongzhen & Yi Liqian 1986: 80).

<sup>35</sup>Some forms seem to be pronounced slightly differently, e.g. *nindi*/нинди was given as /nindey/.

Table 5.142: Interrogatives from Chinese Tatar (Chen Zongzhen & Yi Liqian 1986: 34, 79f., 185), Tatar (Poppe 1963: 81f., 219, 234f. passim), Kazakh (Geng Shimin & Li Zengxiang 1985: 53, 103, 172, 238), and Kyrgyz as spoken in China (Hu Zhenhua 1986: 58, 251)

	Ch. Tatar	Tatar	Kazakh	Kyrgyz
who	kim	kem	kim	kim
what	ni, nerse	nii, nãrsã	ne, nemene	emne
why	nik, nige	nik, nigã	nege, ne yfin	emne ytfyn
how many/much	nitfe, qantfa ~ qantfe	ničã	nefe(w), qanfa	netfe(n), qantfa
how	nitfik	niček	qalaj ~ qandaj	qandaj
what (kind of)	nindij	nindi		
which	qaj, qajsə	qaj, qajsy	qajsə	qajsə
where (to)	qajda, qaj dzer	qajda	qajda	qajda
whence	qajdan	qajdan	qajdan	qajdan
when	qajtjan	qajčan	qajan	qatjan

Table 5.143: Uyghur (Tuohuti Litifu 2012: 367; Mi Haili 1997: 83), and Uzbek interrogatives (Boeschoten 1998: 373; Landmann 2010: 24)

	Uyghur	Kashgar Uyghur	Uzbek (Boeschoten)	Uzbek (Landmann)
who	kim	kim ~ tfim	kim	kim
where	nã	nεε	-	-
how much	nãččã	netftfe	nêčã ~ nêč(ã)tã	necha, nechta (attr.)
what	nemã	nime	nimã	ne, nima
why	nemišqa ~ nemiškã	nimifqa	nãgã, nimãgã, nãčük, nimã üčün	nega, nimaga, nima uchun
when	qačan	qatjan	qačãn	qachon
how	qandaq	qajdaq ~ qandaq	qaláy, qandáy, qanaqa	qani
how much/many	qančã	qantfe	qančã	qancha
where	qeni, qãyãr	qeni, qeje(r)	qani, qayêrdã, qayãtã	qayerda, qatta
which whither	qaysi	qa(j)si	qaysi qayêrgã, qayãqqa	qay(si) qayerga
whence	qãyãrdin		qayêrdã, qayãqdãn	qayerdan

Kazakh *ne yfin* and Kyrgyz *emne ytfyn*, like Uzbek *nimā üčün*, literally mean ‘what for’. Plural forms in Kazakh are formed by reduplication, e.g. *kim kim* ‘who (plural)’ (Geng Shimin & Li Zengxiang 1985: 54). This pattern that is also found in Uzbek, for example, has parallels in the Amdo Sprachbund.

Uyghur *nä* ‘where’, or its dialectal Kashgar form *nee*, is an innovation also found in Eynu *ne* that might be connected to Turkish *nere* and Khalaj *nīrā* ‘where’. Uzbek *nāgā* ‘why’ has cognates in Tatar *nigā* and Kazakh *nege* and in some Siberian languages such as Khakas *noya* or Fuyu *noyo* and is an old dative form. The Uzbek interrogative *nimāgā* ‘why’ has the same basis but is more readily analyzable as the form *nimā* ‘what’ still exists. The dative can also be found in *qayērgā* ‘whither’. Both Eynu and Ili Turki forms are almost completely identical to Uyghur (Table 5.144).

Interrogatives from the Sayan subbranch of Southern Siberian Turkic languages have been collected in Table 5.145. Ragagnin (2011: 94) only mentions the three Dukhan interrogatives *gim* ‘who’, *jū(ü)* ‘what’, and *gae* ‘which’.

**Abakan** is the only subbranch of Siberian Turkic that lacks the special interrogative that might be cognate with Yakut *tuox*. Table 5.146 summarizes all forms available for Khakas, Fuyu, as well as Shor and compares them with Sarig Yughur. Sarig Yughur interrogatives are rather different from other Abakan languages. Altogether there are more forms starting with an *n*-.

Table 5.147 presents data from Chulym and Altai Turkic languages. Altai has a dialectal difference between southern *ne* ‘what’ and northern *d’uy* ~ *čuy* ‘what’ (Baskakov 1958b: 15). For Chulym, Anderson & Harrison (2006) and Harrison & Anderson (2003) have the form *tʃo* for Middle Chulym, but Birjukovich (1997: 493) mentions *nömā* instead, which was given as *nöömä* by Li Yong-Söng et al. (2008). Chalkan furthermore has a verb *t’uvet* ‘to do what’ and the Russian interrogative *qaqoy* ~ *kakoy* ‘what kind of’.

Northern Siberian Turkic interrogatives (Table 5.148) have two resonances, *t*- and *k*-. The latter changed to *χ*- in Yakut. There is no resonance in *n*-, which stands in stark contrast even with several Southern Siberian Turkic languages. Yakut *χanna* and Dolgan *kanna* ‘where’ are amalgamated forms that go back to a locative form with an *-n*- instead of a *-y*-, cf. Sarig Yughur *qay-ta* ~ *qan-ta*. The similarity to Mongolic languages such as Khamnigan Mongol *kaana* or Buryat *xaana* ‘where’ is thus due to chance. Yakut *χas* and Dolgan *kas* ‘how much’ seem to have a cognate in Tuvan *qaš*.

Table 5.144: Ili Turki (Hahn 1991: passim) and Eynu interrogatives (Lee-Smith 1996a: 857; Zhao Xiangru &amp; Aximu 2011: 79f., 316, 338).

	Ili Turki	Eynu (LS)	Eynu (ZA)
who	kim	kim	ki
how many	näččä		netʃtʃe
what	nemä	nime	neme, qaj
where		ne	ne
when	qačan	qatʃan	qatʃan
how much/many	qanča		qantʃe
how	qandaq	qandaq	qandaq
which	qaysı	qajsi	qajsi

Table 5.145: Russian Tuvan (Anderson & Harrison 1999), Dzungar Tuvan (Wu Hongwei 1999: 42, 231), Tofa (Rassadin 1997: 381), and Karagas (Tofa) interrogatives (Castrén 1857b: 23, 163ff.); according to Schönig (1998: 410), the Tuvan form for ‘who’ is *qim*; some variants were excluded

	Tuvan	Dzungar Tuvan	Tofa	Karagas
who	kim	cəm	kum	kəm, kum
what	čüü	dzy-dʒime	čü	tü
to do what			čoon-	
why	čü-ge, čü-den	dzy-ge, dzy-nen	qančza ‘how’	tü-gä, tü-dän, tüneŋ ušun
how many	čeže	dzeze	čehe, čü hure, qaš	teše, tehe
how, which	kandig	candəy	qandyŋ	kandeg
which	kayi ~ kay(i)zi	caj(ə)sə	qajsy	kaja
where	kayda	cajda	qajda	kaida
whither	kayaa ~ kaynaar			kainar
whence	kayiin ~ kayiiirtan	cajlap	qajdan	kajen
when	kažan	cažan		kašan, kähän

5 Survey of the grammars of questions in Northeast Asia

Table 5.146: Khakas (Anderson 1998: 21), Koibal (Khakas) (Castrén 1857b: 23, 163ff.), Fuyu (Hu Zhenhua & Imart 1987: 31), Shor (Donidze 1997: 505), and Sarig Yughur interrogatives (Roos 2000: 87, modified transcription); Shor forms in brackets are from Nevskaja (2000: 294)

	Khakas	Koibal	Fuyu	Shor	Sarig
who	kem	kem, kim	gim	kem	k <sup>h</sup> im
what	nime	nô, nêmä, nime	n <sup>y</sup> em	noo	ni
why	noya	nôdaŋ, nô kerektäŋ, nimedäŋ, nôderga	ново ~ noo		naγo, <b>nati</b>
when	xayžan	kađen	cajan	[qačan]	qa <sup>h</sup> tan
how	xaydi	kaidi ~ kaidi			<b>niyor</b>
how much/ many	ninže, xanža	nemde	ninji		niçi, <b>niçor</b>
where	xayda	kaida	cayda	kajda, [qayde]	q <sup>h</sup> an
which, how	xay, xayzi	kaize, kaizeder	cayzi	kaj(y)	qaysi
whither	xayya, xaydar	kaidar		[qayaya]	qay-ta ~ qan-ta
whence	xaydaŋ	kaidan			qay-tan
what kind of	xayday	kaidak	cadah ~ cadih		<b>niçik</b>

Table 5.147: Middle Chulym (Li Yong-Söng et al. 2008: 44; Anderson & Harrison 2006; Harrison & Anderson 2003), Altai (Baskakov 1997: 183), and Chalkan interrogatives (Erdal et al. 2013: passim). Not all variants listed.

	Chulym (Li)	Chulym	Altai	Chalkan
who	kim		kem	kem
what	nöömä	tʃ <sup>i</sup> o	ne	t <sup>h</sup> u(u), t <sup>h</sup> ü-γ ~ t <sup>h</sup> u-γ/g, ne
what kind of, how			kažy	qaydat, qayde, qaydeet, qayt(a)
which, how	qaydiŋ	qajduŋ	kandyj	qandiŋ, qandu(γ), qanduu
why	qaya	qaja		qay ‘how’
when	qačan ~ qajan	?qajya	kačan	qažan
where	qayda		kajda	qaya, qayda
whither	qaynar	kajnaar		qana(a), qayda
whence	qaydin	kajdin		
how many/much	köpä		kanča	qanža, qanži, qanča, qanči

Table 5.148: Selected Yakut and Dolgan interrogatives (Stachowski &amp; Menz 1998: 423; Stachowski 1993: passim)

	Yakut	Dolgan
who	kim	ki(i)m
what	tuox	tuok ~ tuogu
why (-DAT)	toyo	togo ~ tuogo
how much	töhö	töhö
when (-DAT)	χahan	kahan ~ kagan
how	χaydaχ	kajda(a)k ~ kajtak
how much	χas	kas
which	χaya	kaja
where, whither	χanna	kanna
whence	χantan	kantan
along where		kanan
how much	χahya	kahya(n)
what kind of	χannik	kannyk

## 5.12 Uralic

### 5.12.1 Classification of Uralic

Leaving aside the possible existence of so-called Para-Uralic for which no direct evidence is available, Uralic may be classified as follows (Janhunen 2009: 65).

- (441) Uralic → Samoyedic  
       Finno-Ugric → Mansic  
           Finno-Khantic → Khantic  
               Finno-Permic → Permic  
                   Finno-Volgaic → Mariic  
                       Finno-Saamic → Saamic  
                           Finno-Mordvinic → Mordvinic  
                               Finnic & Para-Finnic

Uralic is usually divided into two main branches, Samoyedic and Finno-Ugric, the latter of which shows strong internal diversity and can be classified into about seven subbranches. However, only the Samoyedic branch (e.g., Janhunen 1977; 1998; Hajdú 1988) will be treated here. Janhunen (1998: 459) mentions two possible classifications of Samoyedic languages which he calls the conventional and the alternative classification. Both classifications share the assumption that Enets and Nenets as well as Selkup and Kamass are relatively closely related, but differ in whether Nganasan and Mator should be granted a separate status or not. As for Enets and Nenets, the focus here will mostly lie on Tundra Nenets and Forest Enets, mostly excluding other dialects. A language called

Yurats probably was a transitional vernacular between Enets and Nenets and will be excluded for lack of data (Janhunen 1998: 457).

### 5.12.2 Question marking in Uralic

Marking strategies for polar questions have been surveyed by Miestamo (2011) for all of Uralic (Table 5.149). In general, Uralic languages form a relatively clear western border of Northeast Asia. Marking with initial, second position or preverbal particles, question affixes, and question word order are all features that set Uralic apart from other languages in Northeast Asia. Some of these features such as word order for marking polar questions rather have affinity with European languages, especially Germanic (§§4.2.1, 5.5.2.1).

Table 5.149: Polar question marking strategies in Uralic (adapted from Miestamo 2011: 8); Int. = Intonation, IP = initial particle, PP = preverbal particle, FP = final particle, 2ndC = second position clitic, WO = word order, AnA = A-not-A

	Int.	IP	PP	FP	Clitic	2ndC	WO	Affix	AnA
Estonian	+	+		+			+		
Finnish				+		+	+		
Veps	+		+		+				
Votic						+			
Central-Southern Saami	+	+							
Northern Saami		+				+			
Skolt Saami	+					+			+
Mari				+					
Erzya Mordvin	+								
Moksha Mordvin	+								
Komi-Zyrian	+					+	+		+
Udmurt	+				+				
Hungarian	+			+	+				
Khanty	+								
Mansi		+			+				
Enets	+							+	
Kamass					+				
Nenets	+			+				+	
Nganasan	+							+	
Selkup			+						

As we will see in this section, not only the marking of polar questions but also the semantic scope of the question markers differentiates Samoyedic, especially northern Samoyedic, from most other languages in this study.

The most complex system of asking questions can be found in *Nganasan*, which has recently been described by Miestamo (2011: 17). It is worth quoting his good summary in full length.

PIs are expressed by the interrogative mood or by intonation alone. The interrogative mood suffixes are different in different tense-aspect categories (they follow all other verb morphology but the person suffixes). In the **present** (aorist), the suffix is *-ηυ/-ηα*, and this suffix replaces the imperfective and perfective aspect suffixes used in the indicative present. However, the aspect suffixes mark aspect only redundantly (and only in the indicative present): the aspect distinction is a lexical one and imperfective and perfective verbs differ in their stems as well (except for a small number of biaspectual stems) – the semantic distinction is thus not lost in the interrogative. In the **preterite**, the interrogative suffix is *-ηυ/-ηα*, and it replaces the preterite suffix used in the indicative. In the **future** expressed with *-sutə*, the final vowel of the verb (the *ə* of the future marker or the vowel of the person suffix) is lengthened if the verb is in final position in the interrogative. The interrogative **iterative** marker is *-kəə*, which differs from the indicative iterative *-kə* by the lengthening of the vowel. The interrogative **future** may also be expressed by *-ntəηυ/-ntəηα*, which is a combination of the progressive aspect suffix *-ntə* and the present interrogative suffix *-ηυ/-ηα*; according to Larisa Leisiö (p.c.), the aorist and future would differ in the progressive interrogative in that the future would contain two instances of the progressive marker, but in actual usage, this repetition often does not happen and the distinction is then not made formally. The interrogative **renarrative** suffix is *-ha* instead of the indicative renarrative *-hamhu*, i.e. the second syllable of the marker is dropped in the interrogative. Other moods do not take interrogative suffixes, although some of them may be used in polar interrogatives. The remote past and the future-in-the-past are used without interrogative marking in questions. The interrogative mood can also be used in content questions. (my boldface)

The same markings are present not only in polar and content questions, but also in alternative questions. The first two sentences are negative questions, present and iterative, showing that the question markers under negation attach to the so-called negative verb – a feature that Uralic shares with Tungusic (e.g., Hölzl 2015a)—rather than the lexical verb itself. Example (444) is an open alternative question in which the second of the two markers attaches to the interrogative verb.

(442) Nganasan

- a. *maaδ təti ńa-mtə ńi-ηi-η heliʃi-ʔ?*  
 why this partner-ACC.2SG NEG-PRS.Q-2SG help-CON  
 ‘Why don’t you help this friend of yours?’
- b. *tij ηonə-ntu’ ńi-kəə-rj’ hourə-ʔ?*  
 2PL self-2PL NEG-IT.Q-2PL fetch.wood-CON  
 ‘Don’t you go to bring your firewood yourselves?’
- c. *ma-kal’i i-ηυ-ru’ maa-ηυ-ru-ʔ?*  
 home-CAR be-PRS.Q-2PL what-PRS.Q-2PL  
 ‘Are you homeless or what?’ (Gusev 2015a: 109, 121)

Question marking in Nganasan is markedly different from all other Uralic languages as well as from most other languages included in this study. Even from a global perspective, it qualifies as one of the more complex interrogative systems. Because of morphophonological alternations the exact form of the question markers is too complicated to be given here in full detail (see Helimski 1998: 489). As one can ascertain from the second part of the open alternative question, content questions also display the same question marking. When the interrogative has no verbal characteristics, the marking attaches to the verb of the clause.

- (443) Nganasan  
*tjn, maa ɔntj-d'i ɔmɔ tuj-ŋu-ruʔ*  
 2PL what so-inf here come-PRS.Q-2PL  
 ‘What have you come here for?’ (Gusev 2015a: 121)

Polar questions in **Forest Enets** have final rising intonation, while in content questions there is a peak on the interrogative (Siegl 2013: 353). Similar to Nganasan, there is a special past tense question marker that appears in both polar and content questions and combines with polar question intonation. Except for the past tense, questions remain unmarked morphosyntactically. No example for an alternative question was found, although the comparison with other Samoyedic languages suggests that they probably exhibit the double marking type. Forest Enets lacks indirect speech and thus has no indirect questions (Siegl 2013: 198).

- (444) Enets (Forest)  
 a. *uu mosra-sa-d*  
 2SG work-PST.Q-2SG  
 ‘Did you work?’ (Siegl 2012: 402)  
 b. *kokod to-sa-d?*  
 whence come-PST.Q-2SG  
 ‘Where did you come from?’ (Siegl 2013: 355)

The past tense interrogative suffix takes the forms *-sa*, *-d'a*, *-t'a*, or *-č'a*, depending on the preceding word (Künnap 1999b: 27). Interestingly, while the answer to a past tense question of course must also be in the past tense, both the tense suffix as well as the position of the agreement marker differ from the question construction.

- (445) Enets (Forest)  
*karaul-xud to-d-udʔ*  
 PN-ABL.SG come-1SG-PST  
 ‘I came from Karaul.’ (Siegl 2012: 404)

According to Siegl (2012: 403), this unusual situation of a tense suffix following an agreement marker is connected with the development of the question suffix.

In the Enets and Nenets languages, a new secondary past tense construction based on the finite verb and a free-standing auxiliary emerged. Later, the free-standing auxiliary merged with the finite verb, resulting in the unusual ordering where tense follows personal endings. Although the reasons for this unusual instance of change, as well as for the prior tense/aspect system preceding this change, await a more thorough investigation and reconstruction, the triggered change resulted in the emergence of a new mood which is only used in questions with general past tense reference.

It may be worth noting that, typologically, the situation is similar to Nganasan. In both languages there is an integration of question marking and tense (or aspect). But there is only one marker in Enets, while there are several in Nganasan, and there is no formal identity of the respective markers.

In **Tundra Nenets** there is a very similar situation to that in Forest Enets. Polar questions display “pitch raising on the penultimate and ultimate syllables, which may make the sentence-final vowel longer.” (Nikolaeva 2014: 267) Polar, content, and alternative questions exhibit the same past tense question marker *-sa* that has a palatalized dialectal variant *-s’a* and changes to *-se* before agreement markers (Nikolaeva 2014: 97). An *s* (*s’*) regularly changes into *c* (*c’*) following consonants (Nikolaeva 2014: 20).

- (446) Nenets (Tundra, Taymyr)  
*ɲawor-ma-n? xarwa-da??*  
 eat-N-DAT want-2PL  
 ‘Do you want to eat?’ (Mus 2015b: 90, from Nenyang)

Interrogatives are either *in situ* or fronted (Nikolaeva 2014: 266).

- (447) Nenets (Tundra)
- a. *xəqman-ca-n<sup>o</sup>?*  
 say.what-PST.Q-2SG  
 ‘What did you say?’
  - b. *pidə<sup>o</sup> ti-m xada-sa-r<sup>o</sup>?*  
 2SG reindeer-ACC kill-PST.Q-2SG>SG.OBJ  
 ‘Did you kill the reindeer?’
  - c. *wera to-sa, n’i-sa?*  
 PN come-PST.Q NEG-PST.Q  
 ‘Did Wera come or not?’ (Nikolaeva 2014: 95, 265, 267)

Forest Nenets has the same question marker found in Tundra Nenets and Forest Enets and presumably exhibits the same semantic scope. Consider an example of a content question in the past tense.

(448) Nenets (Forest)

*kuńana me-sa-n?*

where COP-PST.Q-2SG

‘Where were you?’ (Mikola 2004: 115)

Another way of forming a polar question usually addressed to oneself is the use of a dubitative enclitic. The enclitic can also be found in content questions and marks alternative questions if used twice, and thus has the same semantic scope as the suffix.

(449) Nenets (Tundra)

a. *n’abako-m’i tuut°-bə-ta=w°h?*

elder.sister-1SG come.FUT-COND-DUB

‘(I wonder) will my sister come or not?’

b. *s’axah tuut°ə=w°h?*

when come.fut=DUB

‘When on earth will he come?’

c. *t’uku yəxa yor’a=w°h n’ii=w°h?*

this river deep-DUB NEG-DUB

‘Is this river deep or not?’ (Nikolaeva 2014: 267, 268)

The dubitative enclitic usually has the form =*m°h* but changes to =*w°h* after vowels and to =(°)*h* after *m*. Examples (444) and (449c) of negative alternative questions exhibiting negative auxiliaries as second alternatives follow a construction very similar to several other languages in NEA.

An interesting alternative question with a focus that is not on the verb is the following in which the verb takes the question suffix. The first alternative precedes and the second follows the verb. The second alternative has rising intonation towards the end.

(450) Nenets (Tundra)

*noxa-m xada-sa-n°, t’on’a-m?*

polar.fox-ACC kill-PST.Q-2SG fox-ACC

‘Did you kill a polar fox or a red fox?’

As can be seen, there is only one question marker. Probably, this is the result of ellipsis of the originally reduplicated verb *xada-sa-n°*. The following example, which I reanalyze as an open alternative question, also has this structure.

(451) Nenets (Tundra)

*ti-m xada-sa-n°, ŋan’i ŋəmke-m?*

reindeer-ACC kill-PST.Q-2SG other what-ACC

‘Did you kill a reindeer or what (did you kill) instead?’ (Nikolaeva 2014: 268)

Tundra Nenets has yet another clitic =*t’iq* ~ =*d’iq* absent in eastern dialects that may be found in questions but is not a question marker as such.

The interrogative clitic is used in questions, most typically, in rhetorical questions, but sometimes also information questions. Its function consists in strengthening the interrogative force, roughly in the same way as the ‘on earth’ expression in English (Nikolaeva 2014: 123)

- (452) Nenets (Tundra)  
*xən'ah m'inja-dəm=t'iq?*  
 whither go-1SG=EMPH.Q  
 ‘Where on earth am I going?’ (Nikolaeva 2014: 123)

There are typologically comparable emphatic elements in Chukchi and Yiddish questions.

For the extinct language **Mator**, only two content questions were recorded. Helimski (1997: 164) claims that both exhibit a suffix *-s* possibly related to the past tense question marker in Enets and Nenets. Given the fact, however, that both sentences were translated into the present tense, this seems rather unlikely. **Mator** has been extinct for over 150 years, which is why more information cannot be obtained.

Unlike all other languages in Northeast Asia, **Selkup** has a preverbal polar question particle derived from the interrogative *qaj* ‘what’ (Miestamo 2011: 18).

- (453) Selkup (Taz)  
*tat qaj qən-na-nti?*  
 2SG Q go-CO-2SG  
 ‘Are you leaving?’ (Wagner-Nagy 2015: 149, from Kuznecova)

The interrogative *qaj* possibly has a Turkic origin (see §5.11.3). Content questions remain unmarked. Wagner-Nagy (2015: 142) is not clear whether final rising intonation affects only polar or also content questions.

- (454) Selkup (Taz)  
*tat qum-it-ip qajitqo ašša apsti-s-al?*  
 2SG man-PL-ACC why NEG feed-PST-2SG.O  
 ‘Why did you not give the people any food?’ (Wagner-Nagy 2015: 142, from Kuznecova)

According to Castrén (1855: 111), alternative questions display the marker *kai* in front of each alternative (a feature shared with Ket, §5.13.2), and in negative alternative questions the second alternative has the form *kai aša?* ‘or not?’ (i.e., *qaj* ‘what > Q’, *ašša* ‘NEG’, Wagner-Nagy 2015).

Miestamo (2011: 15) analyzes **Kamass**, extinct since 1989, as having an enclitic polar question marker *=a*. The marker attaches to the verb and does not appear in content questions. Alternative questions are marked twice with the marker *=bV*, like *=a* given with a hyphen but called particle by Künnap (1999b: 35f.). In line with Miestamo’s (2011) analysis, it is treated as an enclitic here. In addition, the example contains a disjunctive *aali* ‘or’, which comes from Russian (Joki 1944: 189).

5 Survey of the grammars of questions in Northeast Asia

(455) Kamass

- a. *tən orōn tərleḡle məlal=a?*  
 2SG hole.LAT skip.GER can.FUT=Q  
 ‘Can you skip the hole?’
- b. *šinde dəyin amna?*  
 who there sit.PRS.3SG  
 ‘Who is sitting there?’
- c. *man amnō’la=bo aali ma’an jila xallola’=bo?*  
 1SG.LOC live.FUT.2PL=Q or home.GEN people.LAT go.FUT.2PL=Q  
 ‘Are you going to live here at my place or are you going home to your own people?’ (Künnap 1999a: 35, 36)

The question marker =bV could have a Turkic origin (§5.11.2), but note that the extinct Kott language, according to Castrén (1858), has two question markers *â* and *bo*, both of which seem to have parallels in Kamass (§5.13.2).

Table 5.150: Summary of question marking in Uralic

	PQ	CQ	AQ
Forest Enets	V-sa ‘PST’	id.	?2x id.
Kamass	V=a	-	2x =bV + aali ‘or’
Mator	?	?-s	?
Nganasan	V-ŋu/-ŋa ‘PRS’, V-hu/-ha ‘PST’, V-sutə ‘FUT’, V-kəə ‘IT’, V-ha ‘RENARR’	id.	2x id.
Selkup	qaj V ‘what>Q’	-	2x qaj V
Tundra Nenets	V-sa ‘PST’, =w’h ‘DUB’	V-sa ‘PST’	2x V-sa ‘PST’, 2x =w’h ‘DUB’

Table 5.150 summarizes marking of polar, content, and alternative questions. Little information on tag or focus questions is available to me, but possibly there is a tag question marker in Nenets that has the form *-xava* ‘is it not so?’ (Miestamo 2011: 16f.).

In general, northern and southern Samoyedic languages have quite distinct question marking strategies. The form and semantic scope of the northern Samoyedic markers set the languages apart from most other languages in Northeast Asia. Table 5.151 gives an overview of two of the question suffixes in northern Samoyedic and their cognates in southern Samoyedic. The Mator suffix *-s-* has tentatively been added, but its exact meaning remains unclear (Helimski 1997: 164).

As shown in §5.11.2, the Southern Siberian Turkic language Khakas has a similar development from a past tense into a question marker (*-žAŋ*) that seems to have been influenced by Samoyedic. Because of the large geographical distance, the Negidal future question marker presumably has no areal connection to Samoyedic (§5.10.2).

Table 5.151: Samoyedic tense markers based on Mikola (2004: 115f.)

PS	Nganasan	Nenets	Enets	Mator	Kamass	Selkup
PRET * <i>-sâ-</i>	<i>-sua-</i> etc.	<i>-sa-</i> Q	<i>-sa-</i> Q	<i>-s- ?Q</i>	?	<i>-s-</i> , <i>-h-</i>
AOR * <i>-ŋ(â)-</i>	<i>-ŋu-</i> Q	<i>-ŋ(a)-</i>	<i>-ŋ(a)-</i>	<i>-ga</i>	<i>-ŷV-</i> , <i>-gV-</i>	<i>-ŋ-</i> , <i>-n-</i> , $\emptyset$

### 5.12.3 Interrogatives in Uralic

For reasons of space only limited aspects of interrogatives in Samoyedic can be presented here. The interested reader is referred to Mus (2009; 2013; 2015b) and references therein, who has given a very detailed description of Samoyedic interrogatives, especially those from the northern languages, and in particular those from Tundra Nenets. Unfortunately, her description lacks a clear historical or morphological analysis.

All Samoyedic languages have a resonance in *k~* (> *x~* in Tundra Nenets), and thus have K-interrogatives. Only some languages have what is called a KIN-interrogative (e.g., Mator *kim*, Forest Nenets *kim'a*). Both features are inherited from Proto-Samoyedic. Janhunen (1977: 15, 62f., 69, 75, 91) reconstructs the following Proto-Samoyedic interrogatives \**ki.m(ə)* ~ \**kj.mä* 'who', \**ku-* 'what, which', \**ku.nä* 'where', \**kä-* 'what, how', \**kä.nə* 'how much', \**mə* 'what', and \**ə.m-* 'what'. Derivations in individual languages, the meaning of the stems, and whether the reconstructed forms are as clearly analyzable as indicated by the hyphens, remain extremely unclear, however. The first three reconstructions share a resonance in \**k~* and thus are probably related historically. In several languages the initial consonant changed to a fricative in some forms such as Forest Enets *sän*, Tundra Nenets *s'an*<sup>o</sup> that are cognates of Nganasan *kanə* and thus derive from \**kä.nə* 'how much'. Generally, most interrogatives seen below can be grouped with one of these reconstructions. The initial *ŋ-* in Tundra Nenets *ŋəmke* and Forest Nenets *ŋami* is prothetic (Janhunen 1998: 466) and the forms are thus derived from \**ə.m-* 'what' (Janhunen 1977: 15). Note that the *ŋ-* only appears in the Central (*ŋamge*) but not the Western (*amge*) and Eastern dialects (*amge*) of Tundra Nenets (Mus 2015a: 93).

Let us now briefly consider the interrogative systems in individual Samoyedic languages, starting with Nganasan (Table 5.152). There is only one resonance in *k-* and only the categories of PERSON (\**k-*) and THING have special forms without this resonance. The interrogative *maa-d'aa* 'why' is derived from *maa* 'what' with the help of what appears to be an allative. A form *sily/сyлy* 'who', borrowed from Nganasan, is attested for Taimyr Pidgin Russian (§5.5.3.3).

In Forest Enets there is also a resonance in *k~*. The interrogatives *obu* 'what', *še* 'who', and *sän* 'how much' do not exhibit this submorpheme, although the latter two historically had an initial \**k* as well.

The interrogatives meaning 'where', 'whither', and 'whence' have separate forms, although the first two share a stem *ku-*, while the last is based on *ko-*. Instead of *obu* 'what', Tundra Enets has the interrogative *mi'* (Künnap 1999a: 5) or *mii'* (Castrén 1855: 97). Forest Enets exhibits an interesting interrogative with the meaning 'which of two' that has its own paradigms shown in Table 5.154 (e.g., *koki-ju?* 'who of us two').

5 Survey of the grammars of questions in Northeast Asia

Table 5.152: Nganasan interrogatives (Helimski 1998: 500f.; Kortt & Simčenko 1985: passim; Castrén 1855: 47, 49, 50, 65, 74); not all variants listed, accents removed

Meaning	Helimski	Kortt & Simčenko	Castrén
who	sili(-ŋuna)	syly	sele
what	maa(-ŋuna)	ma'	maa
why (?-ALL)	maa-d'aa	ma-d'a	maajaarj
when	kanje	kanga	kanarj
how many	kanə	kano'	kana', kanarj
“the how manieth”	kanə-mtu(ə)	kangkoj	kanamtua, kanagüi'
how many times	kan-ü?		kani'
which, what kind of	kuə, kun <sup>j</sup> iə, kured <sup>j</sup> i	kunie, kurodi, kuredi	kunie, kurajee
where	kunu, kun <sup>j</sup> ini	kuninu	kuninu
whence	kun <sup>j</sup> iδ	kunida	kunida
along where		kunimenu	kunimanu
whither	kun <sup>j</sup> i, kundə	kuni'	kuni'aarj, kunijaarj
how	kun <sup>j</sup> i-ʔi'a	id.	id.

Table 5.153: Forest Enets (Siegl 2013: 195ff.; Künnap 1999a: 5, 22, 27, 30, 40) and Enets interrogatives (Castrén 1855: 76, 81, 82, 90, 91); not all variants listed

	Siegl	Künnap	Castrén
who	še	seea	sio, sie
how much	sän		senno
what	obu	obu, abbu	awuo
why (TRSL)	obu-š		
whence	ko-ko-đ	kuhoδ	kuro, kudo, koohoro
where	ku-ni-n	kunje	kokohone
whither	ku-ʔ	kuo'	kuu
along where			kuuno'one
how	kuń	kuń', kud', ku'on	kuuno' kurahaane
when	kuna		kun(n)e
which, what kind of	kursi	kurse	hooke

Table 5.154: Paradigms of the Forest Enets interrogative ‘which of two’ (Siegl 2013: 198)

	DU	PL
1	koki-ju?	=1DU
2	koki-ri?	koki-ra?
3	koki-đi?	koki-đu?

Apart from other Samoyedic languages (e.g., Tundra Nenets *xujum?* ‘which of two’, Mus 2015b: 79), this interrogative has no functional parallel in Northeast Asia, but in Proto-Indo-European *\*k<sup>w</sup>oteros* (§5.5.3).

Nenets interrogatives exhibit two resonances, one in *k-* or *x-*, and another in *s’-* or *š-*. Initial *\*k-* regularly changed to *x-* in Tundra Nenets, but remained stable in Forest Nenets (Hajdú 1988: 4). The initial *s’-* or *š-* likewise goes back to *\*k-* (cf. Janhunen 1977: 62f.). As mentioned before, the initial *η-* is prothetic (Janhunen 1998: 466).

Table 5.155: Tundra Nenets (Nikolaeva 2014: 50, 265, *passim*), Forest Nenets (Mus 2013: *passim*), and Nenets interrogatives according to Castrén (1855: 3, 10, 32, 327); the Tundra Nenets forms in square brackets are from Mus (2013: 2015b); not all variants listed

	Tundra Nenets	Forest Nenets	Castrén
who	xiib’a	kim’a	hübea, hibea etc.
which	xə-n’a-ŋi°	ku-ńa-ŋi	hu-naa-ŋy
where	xə-n’a-na	ku(-ńa)-na	hu-naa-na
whither	xə-n’a-h	ku(-ńa)-ŋ	hu-naa
whence	xə-n’a-d°	ku(-ńa)-t	hu-naa-d
along where	[xə-n’a-mna]		hu-na-mna
how	xə(n)c’er°q	kuše?, kušel	hunder, hunzier etc.
to say what	xəqman-		?ha-maan
what (kind of)	xurka		hurk(k)a
how many	s’an°	šan	saŋooka, sambir
when	s’ax’h	šaxaŋ, šajna, šana	?saha’
what size	[s’aŋkar]	šam’an	?saŋum, saŋuna ‘how long’
what (kind of)	ŋəmke	ŋami	(ŋ)amge(e)
why	ŋəmke	ŋameʔ, ŋamiŋaš	(ŋ)amge(e)jemne

The form meaning ‘when’ is derived from ‘how many’. Tundra Nenets has only one form, whereas Forest Nenets makes a distinction into different forms for ‘what (kind of)’ and ‘why’. The interrogative *ŋəmke* has the irregular accusative plural *ŋəwo* (Nikolaeva 2014: 25) and exhibits a function similar to *xurka*. The two forms are sometimes interchangeable.

(456) Nenets (Tundra)

*xurka/ηəmke l'ekarə-ηe° tara-sa?*

which doctor-ESS needed-PST.Q

'What kind of doctor did he work as?' (Nikolaeva 2014: 261)

Locative demonstratives and interrogatives in Forest Enets show partly parallel paradigms with special morphological markers *-n* 'LOC', *-ʔ* 'ALL', and *-ḍ* 'ABL' that are otherwise only known from postpositions (Table 5.156).

Table 5.156: Demonstrative and interrogative paradigms in Forest Enets (Siegl 2013: 197, 204); modified analysis

	PROX	DIST	INT
LAT	äu(-ʔ)	to-ni-ʔ	ku-ʔ
LOC	äku(-xu)-n	to-ni-n	ku-ni-n
ABL	äku(-xu)-ḍ	to-ni-ḍ	ko-ko-ḍ

While all three stems share the same case markers, there are differences in the formation of the stems that are only insufficiently understood. Siegl (2013: 204) admits that the "spatial deixis system of Forest Enets is far from being clear". However, a comparison with Tundra Nenets sheds some light on the situation.

In Tundra Nenets the suffix *-ηi°* (~ *-(x)°* ~ *-y°*) in the selective interrogative *xə-n'a-ηi°* is an attributive form (Nikolaeva 2014: 52). The locative usually has the form *-xən(°)a*, the dative has the 2nd and 3rd person possessive form *-xəh-*, and the ablative has the form *-xəd°* (Nikolaeva 2014: 62ff.). Apparently, these forms contain an element *-xə* that is missing in the locative interrogatives that simply add the case markers *-na* 'LOC', *-h* 'DAT', and *-d°* 'ABL', but attach to an element *-n'a* instead that has been translated as 'at, by' (Nikolaeva 2014: 50). The prolativ, found in *xə-n'a-mna*, usually has the slightly different form *-mən(°)a(h)*. Apart from the locative forms listed in Table 5.155, Nikolaeva (2014: 50) mentions the shorter forms *xu-na*, *xu-h*, *xu-d°*, and *xu-mna*. This variation can also be seen in Forest Nenets, e.g. *ku(-ña)-na* 'where'. Forest Enets shows a less clear picture, but it can be noted that both the case markers (*-n*, *-ʔ*, *-ḍ*) and stem formations (*-ni*, *-xu*, *-ko*) have parallels in Tundra Nenets (*-na*, *-h*, *-d°*, and *-n'a*, *-xə*, *-ko*). The last of the suffixes can perhaps be found in Tundra Nenets demonstratives such as e.g., *t'u<sup>o</sup>ko-xə-na* 'there', which seems to correspond to Forest Enets *to-ni-n* but has different derivations. The comparison with Nganasan in Table 5.157 illustrates basically the same pattern. Helimski (1998) recorded synchronic variation in Nganasan with (*ku-n'i-ni*) and without the stem extension (*ku-nu*) as well.

The Tundra Nenets interrogative stem *xə-*, mistakenly called an "interrogative prefix" by Wagner-Nagy (2016: 3204f.), fused with the negative verb *n'i-*, resulting in the complex form *xən'a-* 'how not' (Nikolaeva 2014: 281). The interrogatives *xiib'a* '(to be) who' and *ηəmke* '(to be) what' may be either verbal or nominal without requiring any derivation (see §5.4.3 on Yupik and §5.10.3 on Tungusic).

Table 5.157: Paradigms of the locative interrogative in Nganasan and Tundra Nenets

	Nganasan	Tundra Nenets	
where	ku-ni-nu	xə-n'a-na	xu-na
whither	ku-ni-'	xə-n'a-h	xu-h
whence	ku-ni-da	xə-n'a-d°	xu-d°
along where	ku-ni-menu	xə-n'a-mna	xu-mna

## (457) Nenets (Tundra)

- a. (*pidər°*) *xiib'a-n°-s°?*  
 2SG who-2SG-PST.Q  
 'Who were you?'
- b. *xiib'a-h teda səwa?*  
 who-GEN reindeer.3SG good  
 'Whose reindeer is good?' (Nikolaeva 2014: 257, 251)

Full paradigms are not attested but see Mus (2009; 2015b) for a partial list of forms.

In Tundra Nenets there is an interrogative *xəqman-* with the meaning 'to say what', with the verb *man-* 'to say' as a second element (see 447 above). Given the special meaning, one cannot exclude an areal connection to Kolyma Yukaghir *monovod-* with the same meaning that exhibits the verb *mon-* 'to say' as a first part (§5.14.3). The verb for 'to say' was already similar in the respective proto-languages (Nikolaeva 2006: 274), but the mere existence of an interrogative with this specific meaning in NEA is extremely rare and might indicate a contact phenomenon.

The extinct language **Mator** had a resonance in *k-* (e.g., *kijm* 'who', *kumna* 'how many', *kulgu* 'which', *kagan* 'when') and at least one form, *amgan* 'why' (Helimski 1997: passim), without it that might be connected with Tundra Nenets (*ŋ*)*amge* 'what'. As in Nganasan, Enets, and Nenets, the locative forms seem all to be built on a stem *ku-*, but no stem extension can be found, e.g. *ku-na* 'where', *kuŋa* 'whither', *kuj* 'whence'. Mator *kulgu* 'which' could correspond to Tundra Nenets *xurka*.

The **Selkup** interrogative system (Table 5.158) exhibits two resonances in *k-* and *q-*. The form *kuti* ~ *qod* seems to have replaced the original form meaning 'who'. The interrogatives *qaj*, and *kaindek* (and less likely *kuššan* ~ *quzan*) seem to derive from a Turkic source (§5.11.3). According to Castrén (1855: 111), Selkup also has an interrogative *kak* ~ *kaŋ* 'how' that was borrowed from Russian *kak/kak*.

Kamass (Table 5.159) has two resonances in *g-* and *k-*, both of which derive from *\*k-*. The initial *š-* in the interrogative meaning 'who' goes back to *\*k-* as well (Janhunen 1977: 69). The individual forms remain largely obscure synchronically.

In sum, the interrogative systems in Samoyedic display a bewildering diversity of forms that in this study is only overcome by Indo-European and Trans-Himalayan. No interrogative has been fully preserved in all Samoyedic languages, many exhibit idiosyncratic derivations, and only a few forms have a relatively wide distribution (e.g., *\*ku.nă*

5 Survey of the grammars of questions in Northeast Asia

Table 5.158: Selkup interrogatives from different dialects (Wagner-Nagy 2015: 152, passim; Castrén 1855: 111, 113, 126); not all variants listed

	Taz dialect	Ob dialect	Castrén
what	qaj	qaj	kai
why	qajqo		kaitko, kaiŋo
who	kuti	qod	kud, kod
when	kuššan ~ kuššat	quzan	kussai, kušal, kunzei
where	kun ~ kut	kun, kut't'an	kun, kaigan
whither	kuččä	qu, kučēt	ku, kaimd
whence	kuuni	kut'aut, quute	kun, kaigan
how			kaindek
how much			kaana

Table 5.159: Kamass interrogatives (Künnap 1999b: 19, 26, 28; Castrén 1855: 179, 180, 181, 183, 184; cf. Joki 1944: 145)

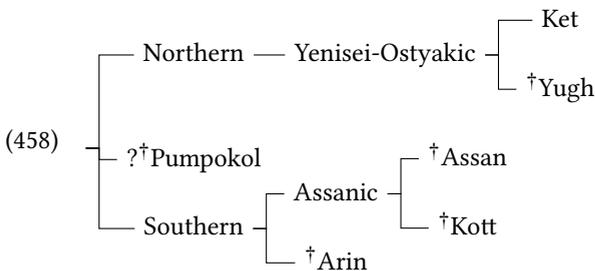
	Künnap	Castrén
what	(ə)mbi(i)	ümbi
why, what for	əmbiile, moo', mo	ümbi ila', ümbi naaman
who	šjindi, šəndi, šində, šəmdə	šimdi
what (else)	šombi	
which, who (of the two)	gi'i'	gidi, kidi
which, what kind	gigə'	
where (to)	gijen	gidigän
where to	giibər	gidibir, gidre
where from	gi'i'	gidigä'
how	kadə', kədə'	kada'
which, what kind	kajet, kəjet	kadet 'how'
when	kaamən	kaaman
how many/much	kümen, gilzi	khümän
which	kümeŋgit	khümäŋit 'the how manieth'

'where'), which either indicates strong language contacts or, what is more likely, perhaps a longer time of separation than the usually accepted 2000 years (e.g., Janhunen 2009). In comparison, Tungusic, which is estimated to be of more or less the same age (e.g., Janhunen 2005), presents a much more coherent picture with many forms found throughout the entire family (§5.10.3). For this reason, the above discussion was not able to give an adequate overview of historical developments, which only an expert in these languages can provide.

## 5.13 Yeniseic

### 5.13.1 Classification of Yeniseic

As we have seen in Chapter 3, the Yeniseic language family differs strongly from most other languages in NEA (e.g., Comrie 1981: 61-66; 2003; Anderson 2003; 2006b; Georg 2008). Today, Ket is the only representative of this language family, but historically there have been more languages, including Yugh (extinct since the 1970s), Kott (extinct since 1850), Assan (extinct since 1800), Arin (extinct since the 1730s), and Pumpokol (extinct since the early 1800s) (Vajda 2009a: 470). Several other languages may have existed but these are almost entirely unknown. This chapter will thus be focusing primarily on Ket, but where possible comparative data will be included from other languages as well, especially Yugh and Kott. There are several attempts at a classification. Georg (2008: 153) proposes the following:



According to Vajda (2009a: 470), Arin can perhaps be classified together with Pumpokol. Both approaches agree in the number of languages as well as in a close relation of Ket and Yugh on the one hand and of Kott and Assan on the other. While Georg classifies Arin with Assan and Kott, Vajda tentatively assumes a connection with Pumpokol. Both approaches are well aware of the somewhat unclear position of Pumpokol. For lack of sufficient information this chapter will exclude Assan, Arin, and Pumpokol. In addition, Vovin et al. (2016), and references therein, have, in my eyes, conclusively shown that at least parts of the Xiongnu confederation in what today is northern China and Mongolia must have spoken a Yeniseic or Para-Yeniseic language (cf. Shimunek et al. 2015), which indicates that, historically, (Pre-)Proto-Yeniseic must have been located much further to the east.

### 5.13.2 Question marking in Yeniseic

Questions in Yeniseic languages have been analyzed by Werner (1995: 155–168), who based his approach on V. A. Moskovoĭ. Unfortunately, his account is rather obscure and lacks a proper analysis of the examples. Where possible, the analysis in this subsection follows Vajda (2004) and Georg (2007).

Polar questions in Ket may take a marker =*u* that usually takes the second position in a sentence, which is a marked difference from most other languages of NEA. Werner

(1995: 159) claims that =*u* is a particle, but wrote it attached to other words with a hyphen. It is reanalyzed as enclitic here.

- (459) Ket  
*tɔːk=ʉ ɛtam?*  
 axe=Q sharp  
 ‘Is the axe sharp?’ (Werner 1995: 159)

Another polar question marker *tām* also converts interrogatives into indefinites, e.g. *tām bila* ‘somehow’ and functions as a disjunction if employed twice.

- (460) Ket  
*bū tām d[u]-i[k]-n-bes?*  
 3SG Q 3M-here-PST-move  
 ‘Did he come?’ (Kotorova & Nefedov 2015: 67)

In negative polar questions the enclitic =*u* attaches to the negator *bən*’ that in this case has the unexplained form *bən’d*. The *d* might be epenthetic.

- (461) Ket  
*tɔːk bən’dʉ ɛtam?*  
 axe NEG.Q sharp  
 ‘Is the axe not sharp?’ (Werner 1995: 159)

Note that the enclitic does not take second position here, perhaps because the negator and question marker were reanalyzed as one element. Possibly, the form has to be reanalyzed as *bən’dʉ* in which the second part might be the unexpected third person singular masculine predicative marker (Stefan Georg p.c. 2016). However, both =*u* and *bən’dʉ* are said to highlight the following instead of the preceding word (Kotorova & Nefedov 2015: 66).

For alternative questions Ket has borrowed the Russian disjunction *ili/или*, used in interrogative and non-interrogative contexts, but also makes use of double marking with the negative polar question marker put before each alternative.

- (462) Ket (Madujka, Kurejka)
- a. *t-a-[i]n-gij, túlet ɔn-am usʃam ili qómat-am?*  
 PREV-THEM-PST-say berry many-PR.3SG EX or few-PR.3SG  
 ‘Say, are the berries many or few?’
- b. *ána qa-du, bándʉ hík.biseb, bándʉ qím.biseb?*  
 who big-PR.3SG.M Q brother Q sister  
 ‘Who is bigger, the brother or the sister?’ (Kotorova & Nefedov 2015: 122, 183)<sup>36</sup>

<sup>36</sup>Many thanks to Stefan Georg (p.c. 2016) for finding these examples and helping with some aspects of their analysis.

In example (462b) an alternative question follows a content question (§4.4). Clearly, the morphosyntactic behavior of the question marker *qaj* in the Uralic language Selkup that appears once before each alternative in alternative questions has an areal connection to Ket *bəndu* (§5.12.2). But while the Selkup marker seems to derive from an interrogative of perhaps Turkic origin (see §5.13.3), this is not the case in Ket. In another example only one marker is present between the two alternatives. In one case a negative alternative question seems to exhibit juxtaposition.

(463) Ket (Kellog)

- a. *ūk huʔn hánuna-da bəndu ūk hiʔp hánuna-du?*  
 2SG.GEN daughter small-PR.3F Q 2SG.GEN son small-PR.3M  
 ‘Is your daughter or your son smaller?’
- b. *ū bín-ku kénas-keʔt, bín-ku bəñ?*  
 2SG self-2SG bright-human self-2SG NEG  
 ‘Are you a real (lit. bright) person or not?’ (Kotorova & Nefedov 2015: 199, 228)

Content questions in Ket are generally unmarked. Interrogatives may be incorporated and thus defocused. Under “object focus” the interrogative *ákùs* ‘what’ takes the form *aj* and under “subject focus” *an* (Vajda 2004: 88).

(464) Ket

- a. *ákùs dš-b-bèt?*  
 what 3F.S-3N.O-do  
 ‘Just what is she making?’
- b. *da-ákùs-[s]-bet?*  
 3F-what-MS-do  
 ‘What is she doing?’
- c. *aj dš-b-bèt?*  
 what 3F.S-3N.O-do  
 ‘Just *what* is it she is making?’
- d. *an kú-[i]n-à?*  
 what 2S-PST-active.event  
 ‘Just *what* happened to you?’ (Vajda 2004: 88)

Both *an* and *aj* are sometimes called question particles (Werner 1995: 156; Kotorova & Nefedov 2015: 66), which clearly must be rejected. For Yugh there seems to be the same mistake (Werner 1997b: 214).

(465) Yugh

- an di'n'e?*  
 what 1S.PST.active.event  
 ‘Just *what* happened to me?’ (Werner 1997b: 225)

## 5 Survey of the grammars of questions in Northeast Asia

Perhaps, *an* has the same function as in Ket as there are also other forms such as *assa* ‘what’ that may also be incorporated and is thus comparable with Ket *ákùs* ‘what’. Historically, Yugh *assa* may go back to \**aksa* (Werner 2004: 157), which makes it even more similar to the Ket interrogative. A cognate of *aj* does not seem to be attested.

(466) Yugh

a. *u assa ku-b-bet’?*

2SG what 2S-3N.O-do

‘What are you doing?’

b. *u k-assa-i’get’?*

2sg 2S-what-do

‘What are you doing?’ (Werner 1997b: 225, 226)

Apart from *an*, Werner (1997b: 214) claims that there are several more question markers, namely *χala* ~ *χara*, *atá*, and *bən*. The status of the first could not be settled,<sup>37</sup> but *atá* most likely is simply an interrogative meaning ‘why’ (§5.13.3), while *bən* is a negator. Werner translates the following sentence with ‘or not’, which seems comparable to example (463b) from Ket above.

(467) Yugh

*dilatkat bil’l’a dənənd’in, bən’?*

children all ?3.PST.3.come NEG

‘Have the children all come or not?’ (Werner 1997b: 225)

For marking polar questions, Yugh had in addition an unspecified intonation pattern (Werner 1997b: 225).

Even less information than for Yugh questions is available for Kott. But apparently, Kott has a final question marker that is analyzed as enclitic here.

(468) Kott

*huš=bo?*

horse=Q

‘(Is it) a horse?’ (Castrén 1858: 153, Werner 1997a: 80)

For lack of further examples the semantic scope of *=bo* remains unclear. Alternative questions seem to take two markers (A=*bo* B=*bo*), although no example was provided by Castrén (1858: 153). Most likely, *=bo*, like the marker *=bV* in the Uralic language Kamass (§5.12.2), has been borrowed from a Turkic source (§5.11.2). Castrén (1858: 154) furthermore mentions the Kott question marker *â*. There is no information on its morphosyntactic behavior or exact function, but it might be connected with the Kamass polar question marker *=a*. Apparently, Russian *li/ли* has also been borrowed. There is no example for a content question from Kott.

<sup>37</sup>A connection with a Mongolian question tag (e.g., Dukhan *hala* ~ *harən*) seems too far-fetched (§5.8.2).

Table 5.160: Summary of question marking in Yeniseic

	PQ	CQ	AQ
Ket	#A=u, #A bóndu, #A tām	-	(bóndu) A bóndu B, ili ‘or’
Yugh	-	-	?
Kott	=bo#, ?â	?	2x =bo#

### 5.13.3 Interrogatives in Yeniseic

The Yeniseic interrogatives strongly differ from those found in other languages of North-east Asia. Especially the large amount of forms meaning ‘who’ and ‘what’ is exceptional. Ket additionally has analyzable forms such as *ásès bi* ‘what kind of thing’ and *ásès ke* ‘what kind of person’ (Vajda 2004: 32). The existence of special female and male forms of the personal interrogatives is unique but has some typological parallels in the Indo-European selective interrogatives (§5.5.3). Ket and Yugh interrogatives usually start with *a*- or with *b*-, which has no clear parallels in NEA, but in Burushaski, for example (Yoshioaka 2012). It may be remembered that this is first and foremost a typological classification and does not necessarily indicate a genetic connection. The interrogative systems in Ket and Yugh are certainly similar to each other and show some direct cognates (e.g., Ket *bísèŋ*, Yugh *bisa<sup>h</sup>:ŋ* ‘where’) and identical categories (e.g., ‘who.SG.F’ vs. ‘who.SG.M’). But there are several striking differences (marked with boldface) that suggest a considerable time of divergence.

Table 5.161: Ket (Vajda 2004: 31, 41f., 88) and Yugh interrogatives (Werner 1997b: 10, 98f., 103, 211, 214, 226); the Ket forms in square brackets are from Georg (2007: 167)

Meaning	Ket	Yugh
who.SG	ánà ~ ánèt	aneit, anet ‘who.SG.M’
who.PL	ánèt-aj, bílānsan	asein, ase:n
what	ák(ù)s, an, aj	assa, an
how many	ánùn, [bilon]	an’ej(a), birej <sup>h</sup> :n, birøn
which, what kind of	ásès, as	aseis, aš’eiš’(i)
when	áskà	aškej
why	áksdiŋt	atá, asesaj
who.SG.F	bésà	asera
who.SG.M	bítsè	aneit, anet
where	bísèŋ	bisa <sup>h</sup> :ŋ
whither	bíltàn, [biles]	bire <sup>h</sup> s, bire <sup>h</sup> :š
whence	bílil, [bili(ŋa)l]	biri:r, birə:r
how	bílà, [bílunon]	birej

5 Survey of the grammars of questions in Northeast Asia

According to Georg (2007: 165), the form *bil̄a-ŋ-s-an* ‘who.PL’ can be analyzed as ‘how-PL-N-PL’ with an unexpected plural marker *-an*, but a development from ‘how’ to ‘who’ is extremely unlikely. According to Vajda (2013: 89), Ket *bi-l-es* and Yugh *bi-r-ε<sup>h</sup>:š* ‘whither’ can historically be analyzed as ‘which-POSS-open.space’. Diachronically, the actual stem thus may not be *bil-* (Georg 2007: 167), but *bi-*.

Table 5.162: Paradigm of the Ket locative interrogative (Georg 2007: 167)

	SG	PL
1	biseŋ-di, biseŋ-am (N)	biseŋ-dan
2	biseŋ-ku	biseŋ-kan
3	biseŋ-du (M), biseŋ-da (F)	biseŋ-aŋ

Interestingly, Werner (1997b: 226) also mentions the Yugh forms *bi-da* ‘where is it/she?’ and *bi-du* ‘where is he?’ that seem to show a gender contrast. This is comparable with the Ket forms *biseŋ-da* and *biseŋ-du* (Table 5.162) that are based on an extended stem (cf. Yugh *bisa<sup>h</sup>:ŋ* ‘where’). Perhaps, Ket *bilon* ‘how many’ is based on the European pattern (e.g., Russian *kak mnogo/как много*), cf. Ket *bil̄a* ‘how’ and *on* ‘many, much’. Note that Yugh, apart from *bir̄on*, has a more transparent form *birej<sup>h</sup>:n*. Ket *ákus* ‘what’ has an abbreviated variant *ák(ù)s* that “must be quite old and stabilized, since the retention of phonetic [k] in the longer variant can only be understood as a remodelling [sic] after the former.” (Georg 2007: 82, fn. 92) This is the basis for *áks-dij̄t(a)* ‘why’, which exhibits an adessive marker (Georg 2007: 166).

For Kott there is an extensive description by Castrén (1858) that has been elaborated on by Werner (1997a). The Kott interrogative system (Table 5.163) also has a resonance in *b~* but only one form starting with *a-* and also has the form *helem* ‘when’ as well as *šena* or *šina* ‘what’ that deviate from this pattern and are perhaps unrelated to the other forms. They do not appear to have been borrowed from any known language.

Reduplication expresses indefinite meaning, e.g. *bili bili* ‘somewhere, everywhere’ (Castrén 1858: 150). The complex interrogative *šena ôjaŋ* ‘why’ is a transparent combination of *šena* ‘what’ and *ôjaŋ* ‘because, for’ (Castrén 1858: 202). Following Vajda (2013: 88), one may identify a stem *bi-* such as in *bi-l-tuŋ*, cognate of Ket *bí-l-tàn* ‘whither’, that goes back to Proto-Yeniseic *\*wi-l-təñ* ‘which-POSS-path’. The exact analysis of most other forms remains uncertain to me as Ket, Kott, and Yugh have a tendency for opaque interrogative systems. Their historical analysis can only be accomplished by an expert of Yeniseic languages.

Non-selective Interrogative pronouns in Yeniseic have extensive paradigms of case marking (Table 5.164, see Werner 1997b: 98 for Yugh; Werner 1997a: 79f. for Kott). Demonstratives show related paradigms (e.g., Werner 1997b: 97, 103).

In sum, Yeniseic interrogative systems deviate strongly from those in all other languages in Northeast Asia. Apart from the formal differences—there are neither KIN- nor K-interrogatives—there are unusual categories such as a gender distinction in the personal interrogatives, incorporation, and a large number of different interrogatives with

Table 5.163: Kott interrogatives (Castrén 1858: 55, 149ff.)

Meaning	Form
which, who	ašix, PL ašig-ân
where	bili
whither	biltuŋ
whence	bilćaŋ
how many/much	bilipei, bilipêi
what kind of	biťäŋ, PL biťäŋ-an
which	bilituiše
what	šena, šina, no PL
why	šena ôjaŋ/uŋô/uŋôjaŋ
when	helem

the meaning ‘who’ and ‘what’.

Table 5.164: Ket singular interrogative paradigms (Werner 1997c: 140)

Meaning	who.MASC	who.FEM	what	who.MASC	who.FEM
ABS	an’a	an’a	akus’	bit’sē	beš’a
GEN	an’a-da	an’a-d(i)	akus’-t	bit’sē-da	beš’a-d(i)
DAT	an’a-daŋa	an’a-diŋa	akus’-tiŋa	bit’sē-daŋa	beš’a-diŋa
BEN	an’a-data	an’a-dita	akus’-tita	bit’sē-data	beš’a-dita
ABL	an’a-daŋal’	an’a-diŋal’	akus’-tiŋal’	bit’sē-daŋal’	beš’a-diŋal’
LOC	-	-	akus’-ka	-	-
PROS	an’a-bes’	an’a-bes’	akus’-bes’	bit’sē-bes’	beš’a-bes’
ADES	an’a-daŋta	an’a-diŋta	akus’-tiŋta	bit’sē-daŋta	beš’a-diŋta
COM	an’a-(γ)as’	an’a-(γ)as’	akus’-as’	bit’sē-γas, bits’as’	beš’a-γas, beš’as
CAR	an’a-(γ)an	an’a-(γ)an	akus’-an’	bit’sē-γan, bits’an’	beš’a-γan, beš’an

#### 5.13.4 Dene-Yeniseian?

If the Dene-Yeniseian hypothesis (Vajda 2010) has a basis in actual fact, the common proto-language must be several thousand years older than Proto-Yeniseic. It is unlikely that question markers remain stable over such long time spans. Similarities would be expected instead in the interrogative system. There have been previous attempts to correlate Yeniseic and Na-Dene interrogatives, notably by Werner (2004: 157ff.), but, in the

absence of clear cognates and sound laws, any comparison must be preliminary at best. For reasons of space and lack of reliable reconstructions, Na-Dene interrogatives cannot be dealt with here. Nevertheless, it is a possibility that the *type* of question marking that is somewhat less prone to changes than the actual question markers shows certain similarities. Given that the Ket question marker is very different from the surrounding languages—Selkup was most likely influenced by Ket—chances are high that it represents a relatively old and possibly stable feature. Na-Dene consists of Eyak, Tlingit, and the Athabaskan languages. Of course, only a very cursory overview can be given here. According to Enrico (2004: 267), Na-Dene languages have a tendency for “clause-initial clitic interrogative markers”. Perhaps, what Enrico has in mind are sentence initial question markers such as in Slavey (see also §4.2).

- (469) Slavey (Mountain)  
*hí golq fehk'é?*  
 Q moose 3.shot  
 ‘Did he shoot a moose?’ (Rice 1989: 1123)

In Western Apache there are both sentence initial and final question markers that may be used independently of each other or combined.

- (470) Apache (Western, San Carlos)  
*(ya') Katie nldzil (né)?*  
 Q PN strong Q  
 ‘Is Katie strong?’ (de Reuse 2006: 57)

However, Eyak, Tlingit, and Navajo all show second position clitics as does Ket.

- (471) Eyak  
*sAsinL=sh?*  
 ?die=Q  
 ‘Did it die?’ (Krauss forthcoming: 648)
- (472) Tlingit  
*lingít=gé x'eeya.áxch?*  
 PN=Q 2SG.understand.it  
 ‘Do you speak Tlingit?’ (Cable 2007: 74, fn. 40)
- (473) Navajo  
*dichin=ish nílí ?*  
 hunger=Q 2SG.POSS  
 ‘Are you hungry?’ (Young & Morgan 1987: 23)

A difference from Ket is the presence of an overt second position question marker in Eyak, Tlingit, and Navajo content questions as well.

- (474) Eyak  
 [*dee k'u.tse'*]=*d*?  
 what meat=Q  
 'What meat?' (Krauss forthcoming: 586)
- (475) Tlingit  
*daa=sá uwajée wutoo.oowú?*  
 what=Q 3PL.think 1PL.bought.it  
 'What did they think we bought?' (Cable 2007: 69)
- (476) Navajo  
*haa=sh yidzaa?*  
 what=Q 3ACC.3NOM.happened  
 'What happened to him?' (Fountain 2008: 33)

In Navajo, the polar (=ish ~ =sh) and content question markers (=shq' ~ =sh) partly overlap in form (Young & Morgan 1987: 23). However, some languages such as Slavey and Western Apache have unmarked content questions just like Ket.

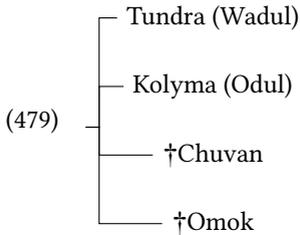
- (477) Slavey (Hare)  
*menj ahət'e?*  
 who 3.COP  
 'Who is it?' (Rice 1989: 1141)
- (478) Apache (Western, San Carlos)  
*hadín dázhó ndzil?*  
 who very strong  
 'Who is very strong?' (de Reuse 2006: 50)

Independent of the question of whether Yeniseic and Na-Dene are genetically related—which cannot, of course, be proven by typology—, Na-Dene shows markedly different question marking than most of NEA, except Ket, parts of Chukotko-Kamchatkan, and some Indo-European languages. Thus, there is a relatively clear boundary between NEA and North America. Furthermore, Dryer (2013l) has shown that polar question marking in the Americas in general is much less uniform than that in NEA. The frequent sentence initial position of interrogatives likewise differentiates Na-Dene and the Americas in general from NEA (Dryer 2013k).

## 5.14 Yukaghiric

### 5.14.1 Classification of Yukaghiric

Today, there are two surviving but endangered, or rather moribund, Yukaghiric languages called Tundra Yukaghir (Wadul) and Kolyma Yukaghir (Odul). Two varieties called Chuvan and Omok are usually included in the list of Yukaghiric languages. Both have already disappeared and are not well recorded (e.g., Anderson 2006e).



According to Nikolaeva (2008), however, “the linguistic status of Chuvan and Omok did not much differ from the status of other varieties of Old Yukaghiric, and therefore referring to them as separate languages within the larger family to the exclusion of other known Yukaghiric idioms is unnecessary”. Old Yukaghiric is a cover term used by Nikolaeva (2008) for those varieties recorded during the 18<sup>th</sup> and 19<sup>th</sup> centuries. Given the limited information on languages other than Tundra and Kolyma Yukaghir this chapter will be concerned primarily with these two modern languages.

### 5.14.2 Question marking in Yukaghiric

**Kolyma Yukaghir** makes a difference between polar and content questions in that only the latter take morphological marking. Polar questions are either expressed with rising intonation or take an enclitic=*duu* that appears twice in alternative questions as well as in negative alternative questions (Nagasaki 2011: 245; Maslova 2003a: 475-478). The semantic difference between polar questions marked with intonation or the enclitic is unknown to me.

(480) Yukaghir (Kolyma)

- a. *omo-s' šoromo-k?*  
good-ATTR person-PRED  
'Is (he) a good person?'
- b. *me-n'oho-j=duu?*  
PRED.FOC-fall-INTR.3SG=Q  
'Did he fall down?'
- c. *igeje čičegej-gen! ad-i=duu šašaqa-daj-m=duu?*  
rope stretch-IMP.3SG strong-INTR.3SG=Q tear-CAUS-TR.S=Q  
'Let the rope stretch! Is it strong (enough), or will he tear it up?'

- d. *kudede=duu oj-l'e=duu?*  
 kill=Q            NEG-be=Q  
 'Have I killed it or not?' (Maslova 2003a: 475-477)

The enclitic also exists in the Turkic languages Yakut and Dolgan, where it has the form =*duo* ~ =*duu* (§5.11.2). The enclitic does not exist in Tundra Yukaghir (Nikolaeva 2006: 150), which also suggests a Turkic origin.

Schiefner (1871) published some material of a variety spoken along the Anadyr that is closely related to Kolyma Yukaghir but possibly has affiliations with Chuvan (Nikolaeva 2006: 28). This variety does not appear to exhibit the enclitic. Instead, polar questions remain unmarked and (negative) alternative questions have a disjunction of Russian origin (Nikolaeva 2006: 101). The tentative analysis roughly follows Maslova (2003a).

- (481) Yukaghir (Kolyma, Anadyr)  
*mot adó kêt'                      alí el=kêt'?*  
 1SG son come.INTR.3SG or NEG=come.INTR.3SG  
 'Did my son come or not?' (Schiefner 1871: 92)

This absence of the enclitic in a variety of Yukaghir spoken further away from Yakut and Dolgan is a further indication that it can be traced back to Turkic. Nagasaki (2011: 254) recorded a tag question that was formed with the help of Russian *da/да* 'yes' attached to a declarative sentence.

According to Maslova (2003b: 66f.) polar questions in **Tundra Yukaghir** are formed with the help of the apparently sentence initial particle *eld'e*. But the particle also appears in content questions.

- (482) Yukaghir (Tundra)  
 a. *eld'e, tide-ŋ                      mit t'ald'ed'uo el=men'-me-k?*  
    well DEM.INV-FOC 1PL ring                      NEG=take-TR-2SG  
    'Well, haven't you taken that ring of ours?'  
 b. *eld'e, neme-le                      men'-me-k?*  
    well what-FOC take-TR-FC  
    'Well, what have you bought?' (Maslova 2001: 48, 42)

A sentence initial question marker indicates a connection with Chukotko-Kamchatkan (§5.3.2). However, as the translation indicates, the word *eld'e* is probably not a real question particle. Neither the exact meaning, nor its origin are discussed by Maslova. Nikolaeva (2006: 154f.) assumes an underlying stem \**el-* that could mean something like 'good', apparently unrelated to the negator *el=* as seen in (482a).

Maslova (2003b: 66f.) mentions two further particles, the dubitative *quolem* (formally similar to interrogatives starting with *quo-*) and hesitant *ejk*. Furthermore, she claims, "if these particles are absent, the verb takes the Negative marker". However, on the same page she gives an example of what appears to be a polar or focus question that neither shows the particles, nor negation.

5 Survey of the grammars of questions in Northeast Asia

- (483) Yukaghir (Tundra)  
*tet-ek Id'ilwej?*  
 2SG-FOC PN  
 'Are you Idilway?' (Maslova 2003b: 67)

Data given by Schmalz (2012) confirms the hypothesis that unmarked polar questions do not have any of the above mentioned particles. Consider the following examples with focus marking on a constituent and on the verb, respectively.

- (484) Yukaghir (Tundra)
- a. *tet-ek werwe-l?*  
 2SG-FOC be.strong-N.SFOC  
 'Are you strong?'
- b. *(mörde(ŋ)) me=möri-mk?*  
 news PRED.FOC=hear-TR.2PL  
 'Have you heard (the news)?' (Schmalz 2012: 69, 71)

Presumably, polar questions can be indicated with intonation only, as is also possible in Kolyma Yukaghir. This suggests a connection with some Chukotko-Kamchatkan languages (§5.3.2).

The proclitic *me=* seen in (484b) can also be found in questions with a denominal verb "to ask for mere confirmation of already known information" (Schmalz 2012: 88). This indicates a functional similarity to tag questions in English.

- (485) Yukaghir (Tundra)  
*me=brigadir-ŋo-d'ek?*  
 PRED.FOC=team.leader-be-INTR.2SG  
 '(You) are the team leader, (aren't you)?' (Schmalz 2012: 88)

Alternative questions also differ from Kolyma Yukaghir in that Tundra Yukaghir uses a disjunctive connective *ejk*, identical to the alleged hesitative *ejk*.

- (486) Yukaghir (Tundra)  
*uo purie-le ejk samnaldarŋ'e-le aptaa-nu-m?*  
 child berry-ACC or mushroom-ACC gather.INCH-DUR-TR.3SG  
 'Is the child picking berries or mushrooms?' (Schmalz 2012: 83)

It is difficult to decide from the limited data whether *ejk* has to be analyzed as a disjunction or as a single question marker, which is a possible marking pattern in some languages. Schmalz presents one instance of yet another possible disjunction, *uuri*, of unknown origin (Nikolaeva 2006: 445).

- (487) Yukaghir (Tundra)  
*tet ile me=čaal'-uon' uuri me=n'aawe-j.*  
 2SG reindeer PRED.FOC=be.bay-INTR.3SG or PRED.FOC=be.white-INTR.3SG  
 'Is your reindeer bay or white?' (Schmalz 2012: 88)

Yukaghiric content questions are more complicated than other question types and involve morphological marking on the verb. In **Kolyma Yukaghir** there is a split between three different paradigms. Special interrogative marking is the default choice, except for so-called intransitive subjects (better called S) and direct objects (better called O), in which case focus marking is employed (Nagasaki 2011: 245). This distribution has certain ergative characteristics, but within focus marking two paradigms exist for transitive (so-called *me*-participle, Tables 5.168 and 5.169 below) and intransitive verbs (*l*-participle, Tables 5.166 and 5.167 below) (Nagasaki 2011: 240). A questioned A (transitive subject) requires no focus or agreement marking.

(488) Yukaghir (Kolyma)

- a. *tet qanin kelu-k?*  
2SG when come-2SG.Q  
'When do you come?' (questioned peripheral argument)
- b. *kin-tek kelu-l?*  
who-PRED.FOC come-PTCP  
'Who came?' (questioned S)
- c. *tet lem-dik ooʒe-t-mo?*  
2SG what-PRED.FOC drink-FUT-PTCP.2SG  
'What will you drink?' (questioned O)
- d. *kin kudede?*  
who kill  
'Who killed (it)?' (questioned A) (Nagasaki 2011: 245, 240)

In Kolyma Yukaghir, interrogatives either stand sentence-initially or remain *in situ* (Maslova 2003a: 481). Sentence-initial position of interrogatives in NEA is rare, but can also be found in Evenki (§5.10.3). Note the additional predicative focus marker *-(le)k* (which has a special form on these two interrogatives) that is included in the case paradigm by Maslova (1997: 459f, 2003a: 88). It appears on nominal predicates as well as on intransitive subjects (S) and direct objects (O) and is thus not restricted to questions (Nagasaki 2011: 227). According to Maslova (1997: 459) it is zero marked on "third person pronouns, proper nouns, and possessive NPs".

Basically the same pattern of content question marking was already in place in the 19th century, as can be seen from the following sentences given by Schiefner (1871) for the variety already encountered above. Again, the tentative analysis tries to follow Maslova (2003a).

(489) Yukaghir (Kolyma, Anadyr)

- a. *kanin kawe-i-ta-je-k?*  
when go-?PFV-FUT-INTR-2SG.Q  
'When will you leave?' (questioned peripheral argument)

5 Survey of the grammars of questions in Northeast Asia

- b. *kɨn-ak*            *kallu-l*    *ta?*  
 who-PRED.FOC come-PTCP there  
 ‘Who came over there?’ (questioned S)
- c. *mɨt lomdak*            *aa-ta-m?*  
 1PL what-PRED.FOC do-FUT-PTCP.1PL  
 ‘What will we do?’ (questioned O)<sup>38</sup>
- d. *kin*    *ólo?*  
 who steal  
 ‘Who stole (it)?’ (questioned A) (Schiefner 1871: 101, 92)

The focused interrogative *kɨn-ak* seems be closer to Tundra *kin-ek* than to Kolyma *kin-tek*. However, this could also be an artifact of the recording.

Content questions in *Tundra Yukaghir* are better understood than polar questions and exhibit a close affinity to those in Kolyma Yukaghir. There are verbal suffixes that “are only used in specific [i.e. content] questions to peripheral constituents” (Maslova 2003b: 20). Matic’s (2014) summary of how content questions are marked can be seen in Table 5.165. Marking of content questions is thus basically identical to Kolyma Yukaghir.

Table 5.165: Content questions in Tundra Yukaghir (Matic 2014: 132, modified)

	S/O	A	Oblique
agreement	S/O focus	-	interrogative
marking on interrogative	focus case	-	-
position of interrogative	preverbal or sentence initial position		

(490) Yukaghir (Tundra)

- a. *neme-lek*    *joorə-t-ook?*  
 what-INST play-FUT-1PL.Q  
 ‘What are we going to play with?’ (questioned peripheral argument)
- b. *kin-ek*    *ewrə-l?*  
 who-FOC come-PTCP  
 ‘Who has arrived?’ (questioned S)
- c. *neme-lə*    *ńeed’i-t-məŋ?*  
 what-FOC tell-FOC-PTCP.2SG  
 ‘What will you tell (us)?’ (questioned O)
- d. *tet-qanə*    *kin*    *tite*    *gitńar*    *wee?*  
 2SG-ACC who so up.to do  
 ‘Who did that to you?’ (questioned A) (Matic 2014: 131f.)

<sup>38</sup>Note that Kolyma Yukaghir has ‘1SG’ *-me* but ‘1PL’ *-l* (Table 5.169).

The focus marker has the form  $-b(\eta) \sim -(ə)k$  (Matić 2014: 131). Maslova (2003b): 8, 52) gives the form as  $-le(\eta) \sim -(e)k$  and again includes it in the case paradigm. According to Schmalz (2012: 55),  $-le(\eta)$  usually attaches to nouns and  $-(e)k$  to pronouns. Interestingly, *kin(-ek)* ‘who’ thus behaves like pronouns and *neme(-le)* ‘what’ like nouns, which is a common cross-linguistic pattern (§4.3). The variant  $-le\eta$  tends to be a focus marker and  $-le$  an accusative (Maslova 2003b: 54). The obligatory combination of focus markers with certain verb forms has a typological parallel in Japonic, where a similar phenomenon is called *kakari musubi* (§5.6.1). Tables 5.168 and 5.169 exclude paradigms for marking of A (transitive subject) as they have almost no special marking; see (488d), (489d), and (490d). In Tundra Yukaghir the third person pronouns take the forms *tud* and *titt*. The verb furthermore remains unmarked except for third person plural  $-\eta u$  (Schmalz 2012: 56).

Table 5.166: Focus marking in intransitive clauses in Tundra Yukaghir (Schmalz 2012: 56); *uu(l)-* ‘to go’

	Verb Focus		Subject Focus (S)	
1SG	met	mer=uu-je- $\eta$	met-ek	uu-l
2SG	tet	mer=uu-je-k	tet-ek	uu-l
3SG	tude.l	mer=uu-j	tude.l	uu-l
1PL	mit	mer=uu-je-li	mit-ek	uu-l
2PL	tit	mer=uu-je-mut	tit-ek	uu-l
3PL	titte.l	mer=uu- $\eta$ i	titte.l	uu- $\eta$ u-l

Table 5.167: Focus marking in intransitive clauses in Kolyma Yukaghir (Maslova 2003a: 140, 144, 234; Nagasaki 2011: 230); *šohie* ‘get lost, disappear’, *amde-* ‘to die’; constructed in analogy to Table 5.166

	Verb Focus		Subject Focus (S)	
1SG	met	m=amde-je- $\emptyset$	met-ek	šohie-l
2SG	tet	m=amde-je-k	tet-ek	šohie-l
3SG	tude.l	m=amde-j	tude.l	šohie-l
1PL	mit	m=amde-j-l'i	mit-ek	šohie-l
2PL	tit	m=amde-j-met	tit-ek	šohie-l
3PL	titte.l	m=amde- $\eta$ i	titte.l	šohie- $\eta$ i-l

The special interrogative verb endings from both languages are collected in Tables 5.170 and 5.171, comparing them with the declarative endings. The suffixes  $-m(e)$  and  $-je$  that sometimes appear in front of the agreement markers express transitivity and intransitivity, respectively (Maslova 2003a: 141). For the most part, the paradigms in Tundra

<sup>39</sup>This suffix takes the form  $-mle$  if following the future marker  $-te$ .

## 5 Survey of the grammars of questions in Northeast Asia

Table 5.168: Focus marking in transitive clauses in Tundra Yukaghir (Schmalz 2012: 56); *ai-* ‘to shoot’

	Verb Focus		Object Focus (O)	
1SG	met	mer=ai-ŋ	met	ai-meŋ
2SG	tet	mer=ai-mek	tet	ai-meŋ
3SG	tude.l	mer=ai-m	tude.l	ai-mele
1PL	mit	mer=ai-j	mit	ai-l
2PL	tit	mer=ai-mk	tit	ai-mk
3PL	titte.l	mer=ai-ŋa	titte.l	ai-ŋu-mle

Table 5.169: Focus marking in transitive clauses in Kolyma Yukaghir (Maslova 2003a: 140, 144; Nagasaki 2011: 221, 230); *juø-* ‘to see, to look at’, *aa-* ‘to make’; constructed in analogy to Table 5.168

	Verb Focus (FUT)		Object Focus (O)	
1SG	met	aa-t-Ø	met	juø-me
2SG	tet	aa-te-mek	tet	juø-me
3SG	tude.l	aa-te-m	tude.l	juø-mele <sup>39</sup>
1PL	mit	aa-te-j	mit	juø-l
2PL	tit	aa-te-met	tit	juø-met
3PL	titte.l	aa-ŋi-te-m	titte.l	juø-ŋi-le

and Kolyma Yukaghir are extremely similar or even identical. One difference is the presence of a first person singular agreement marker *-ŋ* in Tundra Yukaghir that is absent in Kolyma Yukaghir. The same difference can be observed in the transitive verb focus paradigms (Tables 5.168, 5.169). Furthermore, Tundra Yukaghir has a special second plural ending *-mk* in the transitive paradigm instead of the expected *-mut* (also compare Tables 5.168, 5.169).

There is the possibility that interrogative agreement forms in Negidal—most unusual for a Tungusic language—may be traced back to Yukaghiric influence (§5.10.2). Similar to both Kolyma and Tundra Yukaghir, Negidal has special agreement forms for the first person singular *-m* as well as the plural (inclusive) *-p*, and the third person plural remains unmarked. The formal similarity in the singular is accidental, but the typological parallel is unlikely to be due to chance. However, Negidal has the same marking throughout all question types and combines this with other question markers.

### 5.14.3 Interrogatives in Yukaghiric

Nikolaeva (2006) reconstructed several Proto-Yukaghiric interrogatives. The form *\*kin* ‘who’ is very similar to forms with the same meaning in several surrounding languages

Table 5.170: Tundra Yukaghir non-future endings (Maslova 2003b: 18); for the interrogative only future endings are available, showing the additional future suffix *-t(e)*

	TR NONFUT	ITR NONFUT	Q FUT
1SG	-ŋ	-je-ŋ	-te-m
2SG	-me-k	-je-k	-te-k
3SG	-m-∅	-j-∅	-t-∅
1PL	-j	-je-l'i	-t-uok
2PL	-mk	-je-mut	-te-mut
3PL	-ŋa (FUT -ŋu-te-m)	-ŋi-∅ (FUT -ŋu-te-j)	-ŋu-t-∅

Table 5.171: Kolyma Yukaghir non-future endings according to Maslova (2003a: 140); alternations of *j* not shown here include *d'* and *č* (Maslova 2003a: 43); alternative forms in square brackets according to Nagasaki (2011: 228f.)

	TR NONFUT	ITR NONFUT	Q NONFUT
1SG	-∅	-je-∅	-m
2SG	-me-k/[-mi-k]	-je-k	-k
3SG	-m-∅	-j-∅/[-∅]	-∅
1PL	-j	-je-l'i [-j(ii)-li]	-l-ook/[-uok]
2PL	-met	-je-met [-j(e)-met]	-met
3PL	-ŋaa/[-ŋam] (FUT -ŋi-te-m)	-ŋi-∅ (FUT -ŋi-te-j)	-ŋi-∅

(the so-called KIN-interrogatives, Chapters 3 and 6). The interrogatives *\*qa-* ‘which’ and *\*qo-* (> *quo-* in Tundra Yukaghir) ‘where’ must be related, historically. They suggest a connection between the two categories of SELECTION and PLACE, the latter usually being derived from the former. However, as is often the case, a reconstruction of clear-cut interrogative stems is rather questionable. More generally, Yukaghiric exhibits the common *K~* resonance present in many languages of the area (Chapters 3 and 6). Proto-Yukaghir *\*leme* ‘what’ may have started with an *\*n* instead of an *\*l* (Kolyma Yukaghir *leme* ~ *neme*, Tundra Yukaghir *neme*) as did *\*noŋoon* ‘what for’. Table 5.172 gives a more exhaustive list of forms from the two extant Yukaghiric languages. Most forms start with a *q~*, only a few with *n~* (~ *l~*) and *kin* ‘who’ has a special position in both languages. Interestingly, the functional distribution of the resonances *k~*, *n~*, *q~* is almost identical to Turkic languages (§5.11.3). In contrast to what Nikolaeva’s (2006) reconstructions suggest, the two Yukaghiric languages share several very specific interrogatives that can be traced back directly to the proto-language.

A difference can be found in the locative interrogatives, i.e. Kolyma *qon* versus Tundra *qadaa* ‘where’. Additionally, while in Tundra Yukaghir case markers attach directly to the locative interrogative, the case marker replaces the final *-n* in Kolyma Yukaghir.

5 Survey of the grammars of questions in Northeast Asia

Table 5.172: Interrogatives in Kolyma (Nagasaki 2011: 245; Maslova 2003a: 238, 250) and Tundra Yukaghir (Maslova 2003b: 41)

Meaning	Kolyma		Tundra
	Nagasaki	Maslova	Maslova
who	kin	kin	kin
what	leme	leme ~ neme	neme
what for	noonjon	noonjon	
by what	numun		
to say what	monobod-		
which	qadi		qadi
what (verb modifier)			qadinol
how many/much	qamun	qamun	qabu-n ~ qabu-d
where	qon	qo-n	qadaa
to be where		qol-l'e-	
whence	qot	qo-t	qadaa-t
whither	qanjide	qa-njide	
along what route			qadaa-n
how	qodo ~ qode	qodo	quode
to be how		qodo-l'e-	quode-ban-
at what place	qadun	qadoon-	
when	qanin	qanin	qan'in
why	qodit	qodi-t	quodii
how often	qamlid'e		qaml'id'e
to be how	qodimie-	qodimie	
to be how many/much	qamloo-		qamlal

Schmalz (2013: 186, 208), in his otherwise excellent grammar of Tundra Yukaghir, analyzes the initial *q-* as the analyzable interrogative stem for the interrogatives in Yukaghir, which might be too far-fetched. The resonance in *q~*, of course, could indicate an original etymological connection, but similarities with demonstratives are perhaps better analyzed as the result of an additional resonance phenomenon or paradigmatic analogy to the demonstratives (e.g., Diessel 2003, Bickel & Nichols 2007). Schmalz (2013: 186, 208) also mentions some additional interrogatives for Tundra Yukaghir that have not been listed above, such as *quodeband'e* 'what kind of' that he analyzes as *quode* 'how', *pan-* 'to be', and the participle *-je* etc.

# 6 Interrogative constructions in Northeast Asia: A summary

Chapter 5 presented a very detailed description of questions in NEA based on a classification into language families. This chapter has an areal and typological perspective instead. Unfortunately, the information found for almost all languages is insufficient for an exhaustive typology. Usually, only the elicitation from a native speaker, the existence of a large and modern grammar book or of a specialized description of questions offer enough information. Not only is there insufficient information on individual question types and on the semantic scope of markers and interrogatives, but most descriptions also lack adequate information on intonation. This summary follows the same structure as the previous discussion. §6.1 gives an overview of question marking and §6.2 of interrogatives. A set of 12 maps in the style of the *World Atlas of Language Structures* (WALS), based on a sample of 83 languages for which sufficient information was available, is presented in §6.4. However, except for Figure 6.5 (Dryer 2013l,j), there is no equivalent for them in the WALS. §6.3 evaluates the significance of the grammar of questions with an emphasis on language contact.

## 6.1 Question marking

### 6.1.1 Marking strategies

Chapter 4 introduced a four-way typology based on the markedness a comparison of declarative sentences with polar questions. Table 6.1 and the following discussion is based on the sample of 83 languages. The majority of languages belongs to type 4 while types 1 and 2 are not attested. Type 3 is found in Central Siberian Yupik, Korean, Jeju, and perhaps Nganasan, all of which are located in peripheral regions. Tundra Nenets and Forest Enets show a mixture of types 3 and 4 and were thus excluded.

Table 6.1: Marking of polar questions versus declarative sentences in NEA

	unmarked PQ	marked PQ
marked declarative	Type 1: 0	Type 3: 4
unmarked declarative	Type 2: 0	Type 4: 77

Details of the marking of polar questions are given in Figure 6.5. Altogether 45 out of 83 languages (about 54%) have a sentence-final marker as the major question marking strategy. Deviations can mostly be found in peripheral regions such as Amdo, Korea, the Ryūkyūan Islands, Chukotka, and the lower Yenisei. In comparison, only 314 (about 36%) of Dryer's (2013l) global sample of 884 languages had a sentence-final question marker (particle or clitic). If one considers all the languages that have sentence-final question markers, including those with additional marking strategies (Figure 6.6), the figure rises to 62 (about 75%) languages out of 83. This speaks in favor of an extremely strong areal feature of NEA. Dryer's (2013l) map indicates that adjacent areas to the west and southwest indeed show less sentence-final question markers. However, there is no such clear boundary with MSEA in the southeast. sentence-final question particles are generally more common in verb-final languages such as in NEA, but are also common in SVO languages such as in MSEA (Dryer 2013a: 274, 277). Concerning this feature, Dryer (2013l: Chapter Text) discovered "an area within Asia including mainland Southeast Asia and extending west into India and north through China to Japan and eastern Siberia". This study has demonstrated that almost all of NEA shares the feature as well. There is a clear area around Amdo that extends towards the south and encompasses Trans-Himalayan languages from several subbranches; it is characterized by verbal affixes (§5.9.2.2). This forms a clear boundary towards the south (see also Dryer's (2013j)). A marked difference also exists between NEA and North America (§5.13.4).

The marking of content questions is shown in Figure 6.7. Altogether, 41 (about 49%) out of 83 languages have morphosyntactically unmarked content questions. As opposed to polar questions only 15 languages (about 17%) have a sentence-final particle or clitic exclusively, but 13 (about 16%) have a morphosyntactic marker. However, by counting all languages that have sentence-final markers or affixes among other strategies, the figures rise to 27 (ca. 33%) and 18 (ca. 22%), respectively. The first is restricted to the middle part of NEA, stretching from Japan in the east to Xinjiang in the west. Regarding the latter, there are two possible areas: (1) Koreanic and northern Ryūkyūan in the southeast, (2) Yupik, parts of Samoyedic, Yukaghiric, and perhaps Negidal (but not Turkic), in the north.

Information for alternative questions is unavailable for 34 out of 83 languages (Figure 6.8). Of the remaining 49 languages, 21 (about 43%) exhibit the double marking type, exclusively. These are mostly located in the northern half, but excluding Chukotka and Kamchatka. Of the 19 languages with a mixed type, only Plautdiitsch, Yiddish, and Urumqi Han Chinese do not have double marking as one of several marking strategies. The remaining 16 languages are mostly located in the southern half. In sum, 37 (about 76%) out of the 49 languages, at least as one possibility, exhibit the double marking strategy. Selkup and Ket share a unique double marking strategy in which the respective markers appear before each alternative. In all other languages the markers follow the alternatives. This indicates areal convergence of Ket and Selkup as well as a special position of Ket among the languages of NEA (§3.5). Altogether, 17 (about 35%) out of the 49 languages contain a disjunction, which may or may not be accompanied by other question markers (Figure 6.9). These are mostly located in the southern half of NEA, including Korean, Mongolian, Chinese, Russian, Uyghur, Kazakh, and surrounding lan-

guages. Focusing only on those languages that have single marking in alternative questions (Figure 6.10), there are indications for two areas, (1) a clear area in Amdo with possible connections to Xinjiang (Uyghur) and areas to the south, and (2) Yiddish and Ukrainian that share an areal background in Eastern Europe.

For most languages no relevant data was available for focus questions, which is why no map was created. There is sufficient information to conclude that focus questions in Japonic languages as well as Mandarin tend to contain both a focus and a question marker, while in Tungusic, Amuric, and Aleut as well as Old Japanese and Middle Mongol the same question marker as in polar questions is employed, which usually attaches to the verb in polar questions and to the element in focus in focus questions.

Morphosyntactic question markers tend to be extremely short, usually just two or three phonemes long. This indicates their grammatical function as well as relatively high frequency. From genetic and areal perspectives the brevity of the forms represents both an obstacle and a possible pitfall: the shorter a given form, the more likely are chance resemblances. In fact, it is easy to find identical question markers in languages from around the world. But in most cases geographical distance and a lack of interaction clearly show that these similarities must be due to chance, e.g. Amdo Tibetan =*na*, Sibe =*na*, and Ura =*na*. Question markers usually have a simple form that may be represented as (C)V(V), i.e. they consist of a minimum of one vowel phoneme (e.g., Amdo Tibetan ə-v) that is optionally preceded by a consonant. In some instances there is a long vowel or a diphthong (e.g., Hateruma =*naa*, Khalkha =(y)UU, Ulcha =*nɔv*, Yakut and Dolgan =*duo* ~ =*duu*, Kolyma Yukaghir =*duu*). Note that almost all question markers, independent of their morphological status, share this pattern. Only in some rare cases are there question markers that do not conform to this generalization (e.g., Ket *bánda*, Tundra Nenets =*wʰ*, Nganasan V-*sutə*, Alutor *matka*, Koryak *metʹke*, Xunke Oroqen *ǰǰma*). In most cases it may be surmised that the question marker is of a relatively young age and will be subject to phonetic erosion during future developments. In some languages the marker may consist of one consonant only. However, while it is true that in Uyghur, for instance, the marker may have the form -*m*, it still preserves the more conservative variant =*mu* as well. In short, the pattern (C)V(V), although not universal, is an extremely strong tendency for question markers in NEA and perhaps worldwide.

Disjunctions tend to be longer and follow no clear phonotactic pattern (e.g., Atkan Aleut *asxuunulax*, Xunke Oroqen *aaki*, Russian *ili*, Plautdiitsch *öuda*, Yiddish *odər*, Mandarin *háishì*, Kazakh *ælda*, Sarig Yughur *taʰqǐ*, Tundra Yukaghir *ejk*, *uuri*). Question tags can well be very short (e.g., English *eh?*, German *ge?* etc.), but are usually considerably longer and less homogenous (e.g., Mandarin *duì ma?*, *duì-bu-duì?*, Russian *ne pravda li?*, German *nicht wahr?*, *richtig?*, English *don't you?*, *right?*, Plautdiitsch *es nich zöu?*, Sarikoli *na sou-d=o?* etc.). This proves the special position of disjunctions and question tags with respect to the domain of question marking.

### 6.1.2 Semantic scope

A comparison of polar questions with content questions reveals that 61 (about 73%) out of the 83 languages have different marking strategies (Figure 6.11). In a global sample of 50 languages, Hölzl (2015c) found a comparable figure of 73%. There is evidence for one clear area including Japanese, Koreanic, Ainuic, Written Manchu (not shown on the map), Kilen, Ulcha (not shown on the map), Dagur, Khorchin (not shown on the map), and Ōgami, which has the same marking in polar and content questions. These languages furthermore also tend to mark focus and alternative questions in the same way. When looking at only those languages in which polar and content questions are overtly marked differently (Figure 6.12), there are three clear areas: (1) parts of Ryūkyūan, (2) parts of Mongolic and Turkic, as well as perhaps Kolyma Yukaghir, Middle Korean (not shown on the map), and Gyeongsang Korean (not shown on the map), as well as (3) Amuric and Uilta. These results clearly prove Levinson's (2012a: 13) rather dubious implicational universal wrong: "For all languages that have clear interrogative markers, they mark yes-no questions (or polar questions) differently from Wh-questions (or content questions)." A similar implicational universal by Siemund (2001: 1019)—"if a language uses a particle to mark constituent interrogatives, then this language will also allow the use of this particle in polar interrogatives"—had already been disproved by Hölzl (2015c; 2016b: 23).

The comparison of the semantic scope of polar and alternative questions is severely hampered by several problems. For instance, Kazakh as spoken in China has a question marker =MA that appears in both polar and alternative questions. However, the latter additionally exhibit a disjunction *ældə*. If the disjunction is seen as a question marker, then polar questions in Kazakh are marked differently. If, on the other hand, disjunctions are seen as a different functional domain that in some languages combines with question marking, the question marker of polar and alternative questions is the same. This study decided for the latter alternative. However, Kazakh, like many other languages of NEA employs two of the polar question markers in alternative questions. Again, we face two mutually exclusive possibilities, but this time no clear solution to the problem is available. For 34 languages no information is available for alternative questions. Of the remaining 49 languages, when excluding disjunctions and neglecting the difference between single and double marking, 37 (ca. 76%) exhibit the same marking as in polar questions. Seven languages exhibit a mixed type and only five languages (ca. 10%) exhibit different polar and alternative question marking (see Figure 6.8 and Figure 6.13).

A comparison of the semantic maps of all the question marking systems with the help of the conceptual space is possible for only a handful of languages. For reasons of space, Figure 6.1 only shows a selection of four languages. These and the data above suggest that there is a strong dividing line between content questions and polar questions, which in turn show affinities with both focus and alternative questions. §4.2.2 introduced a possible universal that is shown with dashed lines between content, focus, and alternative questions (Content questions are only marked in the same way as focus or alternative questions if polar questions are also marked in the same way.)

The only possible exception to this rule found in NEA is the Ryūkyūan language Miyara from the Japonic language family (Davis & Lau 2015). In this language, if compared

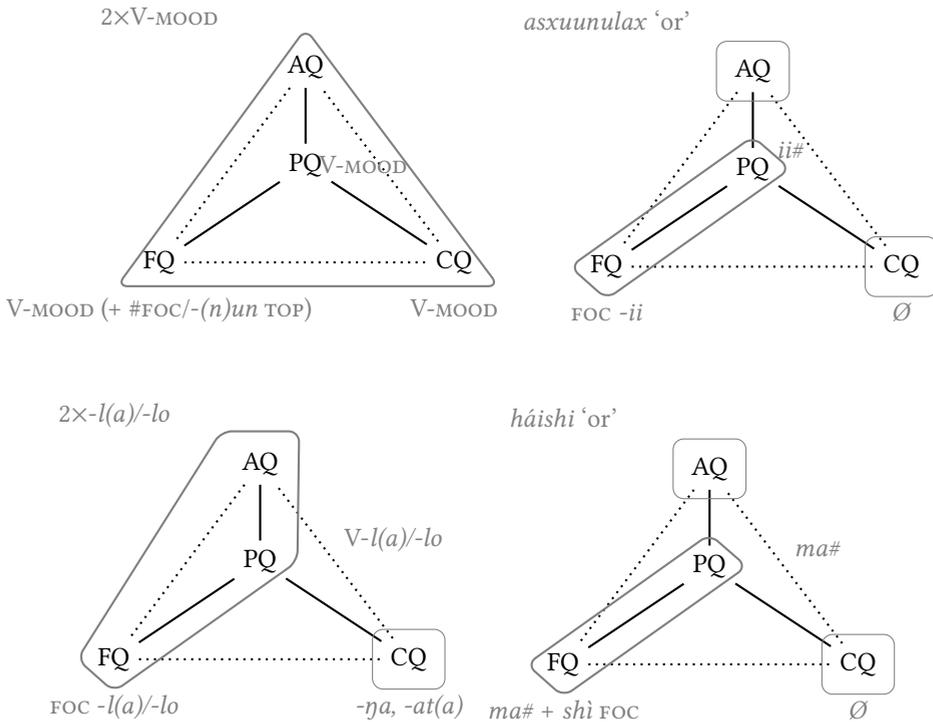


Figure 6.1: Semantic scope of question markers in Korean (top left), Atkan Aleut (top right), Nivkh (bottom left), and Mandarin (bottom right)

with the declarative sentence, both focus as well as content questions lack the indicative marker. But as further specified in §5.6.2, content questions, like declaratives, have falling intonation while polar and focus questions share rising intonation. That the indicative marker is missing results from the fact that both types of questions, content and focus, share a focus marker that is incompatible with the indicative. This is a subtype of the phenomenon usually called *kakari musubi* (focus concord). In the end, Miyara thus most likely presents no exception to the universal. The second universal (Focus and alternative questions can only be marked in the same way if polar questions are also marked in the same way.) seems to hold for Northeast Asia as well. Of course, both universals (or tendencies) can be unified into one form: Focus, alternative, and content questions can only be marked in the same way if polar questions are also marked in the same way.

### 6.1.3 Interaction of functional domains

§4.2.3 identified the following possible interactions of functional domains (see also Hölzl 2016b: 24): (1) grammaticalization, (2) combination, (3) fusion, (4) interaction (split types). Tables Table 6.2 to Table 6.5 give a list of all instances of these interactions in NEA.

6 Interrogative constructions in Northeast Asia: A summary

Table 6.2: Grammaticalization of question markers in NEA (1)

Language	Type	Form	Source Domain
Alutor	PQ	#matka	?INT
Koryak	PQ	#met'ke	?INT
Ukrainian	PQ	#čy	INT 'how'
Tocharian A	?PQ	aśsi	adposition + INT
Sogdian	PQ	(ə)ču-t(i)	INT 'what-COMP'
German	AQ	kataar(-əti)	INT 'which-COMP'
	TQ	was, wie	INT 'what', 'how'
Selkup	TQ	oder	OR
	PQ	qaj V	INT 'what'
Japanese	PQ	tte	QUOT
	PQ	no#	NMLZ < ?GEN
Ainu	PQ	ruwe#, hawe#, siri#	NMLZ
Ōgami	CQ	-εε	?NMLZ
	PQ	=tu#	?FOC
	CQ	=ka#	?FOC
Khorchin Mongolic	PQ	fii	NEG
	CQ	Khalkha be#, Buryat be# ~ =b, Khamnigan Mongol bei#, Oirat =w ~ =b, Shira Yughur bə#	COP
Turkic	CQ	Salar V-i, Tuvan V-Il, Tofa V-(u)l, Dolgan V-ij, Yakut V-(n)ij	DEM > ?COP
	PQ	e.g., Turkish =ml	NEG > Q
Mandarin (dialects)	PQ	ma#	?NEG (NAQ)
	PQ	bu#, mei#	NEG (NAQ)
Amdo Tibetan	PQ	=na	?NEG (NAQ)
Nġanasan	PQ, CQ, AQ	PRS V-ŋu/-ŋa, PST V-hu/-ha, FUT V-sutə, IT V-kəə, RENARR V-ha	TAME
Chukchi, Kerek	CQ	IMP-V	IMP
Forest Enets	PQ, CQ, ?	PST V-sa	tense
Tundra Nenets	PQ, CQ, AQ	PST V-sa	tense

The origin of most question markers is obscure. Several somewhat unclear cases discussed in Chapter 5 such as the Ryūkyūan (=na(a) and variants) or the Yakut and Dolgan question markers (=duo ~ =duu) that could be related to interrogatives meaning ‘what’ were omitted. If Ōgami =ka and =tu indeed derive from focus markers, this is most likely also true for several other Ryūkyūan languages (e.g., Shuri, Tsuken, Tarama, Ikema, Irabu, §5.6.2). Mandarin dialects have also not been listed separately. There are several possible instances of shared grammaticalization such as the development from nominalization to question markers in the Japanese archipelago (Ainuic, Japonic) as well as the development of content question markers from copulas in several Mongolic and Turkic languages (see §5.8.2 and §5.11.2).

Table 6.3: Combination of question markers with other functional domains in NEA (2)

Language	Type	Form	Meaning
Japanese	?FQ	ka# + wa	TOP
Korean	?FQ	V-MOOD + -(n)un	TOP
Wutun	?FQ	=mu#/=a# + -ha	TOP
Yuwan	FQ	V-ui + =du	FOC
Shuri	FQ	=ji# + =du	FOC
	FQ	V-ra + =ga	FOC
Ikema	FQ	=na# + =du	FOC
Irabu	PQ	=ru# + =ru	FOC
	CQ	=ga# + =ga	FOC
	FQ	2x =ru + =ru	FOC
Miyara	FQ	- + lack of IND -N + =du	FOC
Khalkha	FQ	=(y)UU# + intonation	FOC
Mandarin	FQ	ma# + shì	FOC
	AQ	A ne B + háishì	OR
Urumqi Hui Mandarin	AQ	A n <sub>i</sub> <sup>44</sup> B + xε <sup>24</sup> s <sub>i</sub> <sup>21</sup>	OR
Hezhou Chinese	AQ	X n <sub>i</sub> <sup>3</sup> , Y, [haishi] Z etc.	OR
Xunke Oroqen	AQ	2x =jA + aaki	OR
Kilen	AQ	=a# + xəçi	OR
Kazakh	AQ	2x =MA, 2SG V-MI + ælde	OR
Uyghur	AQ	2x =mu, NPST V-m + yaki	OR
Sarig Yughur	AQ	2x =Mi, =mu# ~ V-m + ta <sup>h</sup> qï	OR

Of course, the exact mechanisms and processes involved in these instances of grammaticalization need additional investigation. Table 6.3 excludes negation and interrogatives. As can be seen, question marking most commonly combines with focus marking and disjunctions. Three patterns of fusion have been listed in Table 6.4.

## 6 Interrogative constructions in Northeast Asia: A summary

Table 6.4: Fusion of question markers with other functional domains in NEA (3)

Language	Type	Form	Meaning
Ainu	PQ	an	COP
Shuri	NPQ	ʔa-ran-i	COP
Shirongolic (Mongolic)	PQ	Bonan wu(u), (m)bu, Kangjia vᠬ, mbᠬ, Santa wu, Mangghuer beinu	COP
Shirongolic (Mongolic)	PQ	V-(TAM.)Q	TENSE
Amur Nivkh	FQ	FOC=l(a)/=lo	FOC
East Sakhalin Nivkh	FQ	FOC=l(a)/=lu	FOC
Atkan Aleut	FQ	FOC=ii	FOC
Russian	FQ	#FOC=li	FOC
Even, (Khamnigan) Evenki, Udihe	FQ	FOC=KU	FOC
Udihe	FQ	FOC=nA	FOC

As can be seen from the entries with question markers in Table 6.5, a large number of descriptions fails to mention the criteria for distinguishing between different question markers. At least for Manchu it could be shown that it depends in part on clause type (see §5.7.2). There is a wide variety of different criteria, but many, like question marking itself, are verbal categories (e.g., TAME, agreement, polarity, clause type).

### 6.1.4 Borrowing

Table 6.6 gives a list of borrowed question markers in NEA. Some cases are not absolutely clear. See §3.1 for the methodology of establishing whether a question marker has actually been borrowed. Most cases require an additional evaluation and elaboration by experts of the individual languages.

One of the most widespread markers is Mandarin *ba* 吧. Probably due to its special semantics (§5.9.2.1), it is far more likely to be borrowed than a more neutral question marker such as Russian *li*/ли (mostly used in the written language), and in fact it can be found in many languages of China, from Xinjiang to Manchuria. Some very unclear cases were excluded.

In a few cases it is more likely that similarities are due to chance. Gothic, for example, has a second position question marker *=u* as well as a sentence initial question marker *an*. At a first glance, these are surprisingly similar to Ket second position *=u* and sentence initial *an* ‘what’, but the large distance in both time and space makes a comparison more than doubtful.

Table 6.5: Split types of polar and content question marking found in NEA (4)

Language	Type	Form	Criterion
CSY	PQ, CQ	V-MOOD+AGR.Q	person, number
NSY	PQ, CQ	V-MOOD+AGR.Q	person, number
Sirenikski	PQ, CQ	V-MOOD+AGR.Q	person, number
Kazakh	PQ	=MA#, 2SG V-MI	person, number
Hatoma	CQ	PST -, ATTR + =wa#, =ja# (non-verbal)	clause type
Sonai	CQ	=ga, =ja(a) (non-verbal) + =ba SEL, + =du FOC	clause type
Gyeongsang	PQ, CQ	PQ -na, COP -ka, CQ -no, COP -ko	clause type
Jeju	PQ, CQ	V-MOOD	politeness
Korean	PQ, CQ	V-MOOD	politeness
Hezhou Chinese	PQ	ma <sup>3</sup> , la <sup>3</sup> , ni <sup>3</sup> mu <sup>3</sup>	politeness, polarity
	CQ	za <sup>3</sup> , ni <sup>3</sup> , ni <sup>3</sup> za <sup>3</sup>	politeness, semantic category, gender of speaker
Japanese	PQ	-kana(a) vs. -kashira etc.	gender of speaker
Kolyma Yukaghir	CQ	FOC.CASE, S/OAGR.Q	grammatical relations
Tundra Yukaghir	CQ	FOC.CASE, S/OAGR.Q	grammatical relations
Ōgami	PQ	=ka#, ?=tu, PST, COP, STAT V-εε	tense, clause type, aktionsart
Wutun	PQ	PFV, RES =mu#, IPFV, PROGR =a#	tense, aspect
Nganasan	PQ, CQ	PRS V-ŋu/-ŋa, PST V-hu/-ha, FUT V-sutə, IT V-kəə, RENARR V-ha	tense, aspect
Uyghur	PQ	=mu#, NPST V-m	tense
Forest Enets	PQ	PST V-sa	tense
Negidal	PQ	=Kʊʊ# + FUT V-MOOD+AGR.Q, =i#	tense, person, ?
Shuri	PQ	IND V-mi, NEG V-i, PST V-ti, =naa#	polarity, tense, ?
Written Manchu	PQ, CQ	COP =o, =ni, =nio, =nA, =nu	clause type, ?
Masana Okinoerabu	PQ	IND -ŋ, -Ø + =ŋja# + PST -ti instead of -ta, =na#	tense, ?
Tundra Nenets	PQ	PST V-sa, DUB =w'h	tense, ?
Amdo Tibetan	PQ	ə-v, =na#, =ni#	?
Oirat	PQ	=(y)UU# ~ =ii#	?
Bonan	PQ	V-(TAM.)Q, V-si	?
Hezhou	PQ	ma#, la#, (ni)mu#	?
Xunke Oroqen	PQ	=jA#, =ɔɔ#	?
Udihe	PQ	=Ku#, =nA#	?
Kilen	PQ	=nə#, =a#, =ma#	?
Salar	PQ	=mU#, =U#	?
Sarig Yughur	PQ	=Mi#, =mu# ~ V-m	?
A Nivkh	PQ	V=l(a)/=lo	?
ES Nivkh	PQ	V=l(a)/=lu	?
A Nivkh	CQ	=ŋa, =at(a)	?
ES Nivkh	CQ	=ŋa, =ŋu, =ara	?

6 Interrogative constructions in Northeast Asia: A summary

Table 6.6: Possible instances of borrowing and loan translations of question markers in NEA. When several dialects have a given form, only one was mentioned

	Type		
Yakut, Dolgan =duo, =duu	PQ, AQ	?→	Kolyma Yukaghir =duu
Mandarin ba 吧	PQ	→	Mangghuer ba, Bonan ba, Kangjia ba, Santa ba, Kilen ba, Sanjiazi Manchu ba, Sibe ba, Salar ba, Khorchin ba(a), ?Oroqen baa ~ bæə, Shineken Buryat baa, Solon baa, Dagur baa etc.
Mandarin ma 吗	PQ	→	Bonan ma, Kilen ma, Yibuqi Manchu ma
Mandarin háishi 还是	AQ	→	Yibuqi Manchu xəʃi, Kilen, Oroqen hafi
Russian ili	AQ	→	Ket ili, Chulym ili, ?Kamass aali, ?Anadyr Yukaghir ali
Nivkh -ŋa, =ii	CQ, PQ	→	Uilta =KA, ?=(y)i
Khamnigan Mongol bei	CQ	→	Khamnigan Evenki bei
Russian li/ли	PQ	→	Khamnigan Mongol =li (AQ)
Koreanic, e.g. Middle Korean -ni, -nyo, -(k)o, ?-nja	all	→	Jurchenic, e.g. Manchu =ni, =nio, =o, ?=nA etc.
Mongolic 'cop>Q' (*büi)	CQ	→	Salar -i, Tuvan -Il, Dukhan -Il, Yakut -(n)ij, Dolgan -ij
Mongolian =(y)UU ~ y.ii	PQ, AQ	→	Salar -u ~ -o, Oroqen =ɔɔ, Ongkor Solon =uu ~ =ii, ?Even =i, ?Negidal =i, ?Uilta =(y)i
Dagur =yee	PQ, AQ	→	Oroqen =jE, ?Sibe =jə etc.
Udihe =nu, (?=nA)	PQ, FQ, AQ	→	Kilen =nə
Mongolian youm=aa	PQ	?→	Oroqen =jOOmAA
Khorchin jimεε	CQ	?→	Solon yeeme
Buryat =gü	PQ, AQ	?→	Solon =gi
Ewenic (Tungusic) *=Ku	PQ, AQ	?→	Buryat =gü, Khamnigan Mongol =gu
Uyghur =mu	PQ	→	Salar, Sarig Yughur, Hezhou, Tangwang, Wutun =mu
Uzbek =mi	PQ	→	Tajik =mi
Yukaghir AGR.Q	CQ	?→	Negidal (PQ, CQ, AQ) AGR.Q
Ukrainian čy etc.	PQ, AQ	→	Yiddish ci
Burushaski or Dardic =a	PQ	?→	Sarikoli =o
Kamass =a	PQ	?→	Kott â
(Turkic ?→) Kamass =bV	AQ	?→	Kott =bo
Hezhou la <sup>3</sup> 啦	PQ	→	Santa -la
Old Japanese ya	PQ	?→	Ainu ya
Xining Mandarin le <sup>53</sup> 呢	CQ	→	Kangjia le

## 6.2 Interrogatives

### 6.2.1 Formal properties

This study has emphasized on formal properties of interrogatives such as the overall shape and the initial sounds. Several language families in NEA exhibit a KIN-interrogative: the interrogative meaning ‘who’ in a given language has the form KIN (velar or uvular plosive or fricative, (high) vowel (short or long), (apical) nasal), followed by an optional final vowel, e.g. Turkish *kim*, Forest Nenets *kim’a*, Aleut *kiin* etc. 30 (ca. 36%) out of the 83 languages exhibit KIN-interrogatives (Figure 6.14). The phenomenon can be traced back over considerable time-spans to several of the proto-languages of NEA (Table 6.7). In some instances (indicated with a question mark), the similarity most likely is due to pure chance. Itelmen *k’e*, for example, superficially resembles the other forms, but most likely derives from PCK *\*mikæ* (Fortescue 2005).

Table 6.7: KIN-interrogatives in NEA

Language	Form	Source
?Ainuic	*gu(n)na	Vovin 1993
Atkan Aleut	kiin	Bergsland 1997
Eskimo (ABS.SG)	*ki-na	Fortescue et al. 2010
?Indo-European (NOM.SG)	*k <sup>w</sup> í-s ~ *k <sup>w</sup> ó-s	Mallory & Adams 2006
?Itelmen	k’e < ?*mikæ	Fortescue 2005
Mongolic	*ke-n	Janhunen 2003d
Samoyedic	*kim(3) ~ *kjmä	Janhunen 1977
Turkic	*kim, Oghur *käm	Schönig 1999
Uralic	*ki ~ *ke	Nikolaeva 2006
Yukaghiric	*kin	Nikolaeva 2006

In *some* of the remaining instances the similarity could in fact indicate long distance relationships. However, these limited data cannot, of course, proof any valid genetic unity as was assumed by Greenberg (2000: 217-224). Nevertheless, their similarity as well as the fact that interrogatives meaning ‘who’ appear to be especially conservative, may suggest certain directions that deserve further investigation. The KIN-interrogative seems especially promising due to its wide distribution. Related phenomena are often restricted to only a few language families, such as Turkic *\*qay-* (e.g., Uyghur *qay-*, Khakas *xay-*) and Tungusic *\*Kai* ‘what, which’ (e.g., Alchuka *kai-*, Nanai *xai*). Such cases can often be more readily explained by language contact or chance.

A phenomenon similar to the well-known m-T-pronouns (e.g., Italian *mi*, *ti*, Nichols & Peterson 2013) are so-called K-interrogatives: more than two interrogatives in a given language start with a velar or uvular plosive or fricative (Figure 6.15), e.g. Nanai *xai* ‘what’, *xado* ‘how many’, *xooni* ‘how’, Uyghur *qaysi* ‘which’, *qačan* ‘when’, *qandaq* ‘how’ etc. The consonant must be identical in the different forms. Altogether 39 (ca. 47%)

out of 83 languages exhibit K-interrogatives, which speaks in favor of a very strong areal feature. While more research is necessary to establish their full geographical extent around the globe, at least parts of Eurasia share the phenomenon, e.g. Italian *chi* ‘who’, *che* ‘what’, *quale* ‘which’, all of which start with [k] (my knowledge), or Bengali *ke* ‘who’, *ki* ‘what’, *kon* ‘which’ (Thompson 2012: 202) (see §5.5.3.1). Of course, there are also languages outside of Eurasia with K-interrogatives, but a comprehensive treatment requires a large cross-linguistic sample. Greenberg’s (2000: 217-224) investigation of the alleged “Eurasianic” interrogative starting with *k*- overlaps with my notions of KIN- and K-interrogatives but is fundamentally different. My categories are first and foremost typological in nature and K-interrogatives are only accepted for a given language if at least three interrogatives share the same initial consonant. Middle Korean, for example, which apparently has only one interrogative starting with *h*- does not fulfill this criterion and therefore has no K-interrogatives. Greenberg (2000), on the other hand, merely assumes that the individual forms are all related to each other but does not follow any accepted methodology. Greenberg (2000) furthermore does not clearly differentiate between interrogatives with different meanings but treats them as one category. However, as shown in Figure 6.16, personal interrogatives in 43 (52%) out of 83 languages do not share the same initial consonant and thus should be kept separate. For instance, in Yukaghiric and several Turkic languages the personal interrogative starts with *k*-, but more peripheral interrogatives start with *q*- instead. This phenomenon is not restricted to NEA, but can also be found in other languages, such as the Dravidian language Kurux (Kobayashi & Turkey 2017: 91). In this language, all interrogatives except *ne*: ‘who’ begin with an *e*~. This result indicates that personal interrogatives (and the category PERSON in general), have a very special position and most likely are more stable than most other interrogatives. Tungusic, for example, has the interrogatives \**ŋüi* ‘who’, \**ja*- ‘what’, and a larger group with a resonance \**K*~. Given that the second and at least some of those interrogatives starting with \**K*~ have most likely a Mongolic origin, the interrogative \**ŋüi* could represent an older layer of the interrogative system. In general, the distribution of KIN- and K-interrogatives overlaps with m-T-pronouns and front rounded vowels, which could indicate an old dispersal of languages in Eurasia that may have had its origin in southern NEA (Nichols 2010).

### 6.2.2 Semantic scope

Unfortunately, not much can be said about the semantic scope of interrogatives in NEA. Most descriptions are extremely vague about the exact meaning of interrogatives, which is why no absolute numbers can be given here, but relatively clear cases of polysemous forms have been collected in Table 6.8. Some polysemies are very frequent (e.g., QUANTITY MASS = COUNT, MANNER = REASON), while others are extremely rare. PERSON = THING can only be found in Tocharian B *kuse*, *mäksu*, and perhaps Ainu *ne*- or Mongolic \**ke*-.

The semantic scope of interrogatives gives a specific pattern for every language. Figure 6.2 illustrates this with the help of several Mandarin interrogatives. The largest category encompassing THING, ACTIVITY, REASON, and TIME is formed by *shénme* ‘what’ and

its derivations. In Mandarin dialects there are some deviations from this pattern. For example, *nǎ-* ‘which’, in the form *nǎ-yi-ge* ‘which-one-CLF’, has expanded its scope to include the category of PERSON as well (§5.9.2.1).

Table 6.8: Polysemous interrogatives in NEA

Scope	Example
ACTIVITY = REASON KIND = MANNER	Mandarin <i>gàn shénme</i> ~ <i>gànmá</i>
MANNER = REASON SELECTION = MANNER	Mongolian <i>yamer</i> Mandarin <i>zěnme</i>
SELECTION = PERSON THING = REASON	Tuvan <i>kandig</i> Mandarin dialects <i>nǎ-(yi)-ge</i>
THING = SELECTION PLACE = DIRECTION	Sogdian (ə)čū Manchu <i>ya</i> English <i>where</i>
PERSON = THING QUANTITY MASS = COUNT	Tocharian B <i>k<sub>u</sub>se, mäksu</i> Kolyma <i>Yukaghir qamun</i>

As in this example, innovative interrogative systems are usually based on the categories of THING and SELECTION (e.g., Cysouw 2007). In principle, the pattern can be given for every language for which sufficient information is available. For reasons of space, however, this cannot be accomplished here for all the languages of NEA. However, as shown in Chapter 5, the conceptual space clearly is able to capture the semantic scope and diachrony of most interrogatives in NEA.

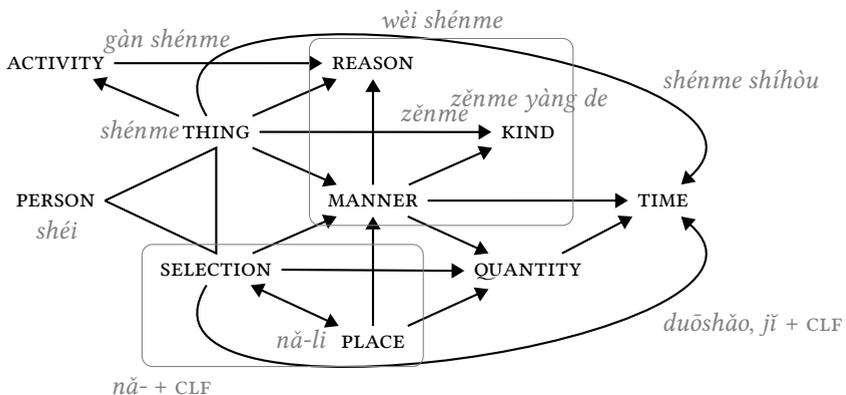


Figure 6.2: Semantic scope of several Mandarin interrogatives

The semantic scope of locative interrogatives can be shown with an additional conceptual space. Figure 6.3 illustrates this with the help of Mandarin data. In Mandarin, all

three categories are marked with *nǎ-li* or its variants, which is, depending on the construction, combined with verbs or prepositions that derive from verbs. However, there are other systems with either identical forms for more than one category and systems with synchronically opaque formations. Some languages for which sufficient information was available are compared in Table 6.9. The individual forms may either be related to each other or not (e.g., Ukrainian). There are several interrogatives that have a scope covering two of the categories (e.g., Japanese, English, Manchu, Mandarin). LOCATION appears to have a tendency to be the unmarked member of the group and often serves as a basis for derivations (e.g., Buryat, English, German, Japanese, Mandarin).

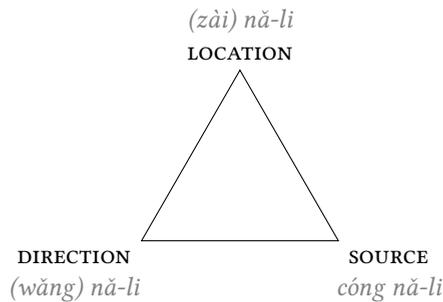


Figure 6.3: Semantic scope of simplified Mandarin locative interrogatives

Many languages in NEA distinguish the three different categories by means of case marking or adpositions (e.g., Mandarin, Khakas, Evenki, Nenets, Kolyemal, Amdo Tibetan, Central Siberian Yupik). However, in some instances not all three forms are based on the same stem (e.g., Buryat, Kolyma Yukaghir).

Information from grammar books is usually insufficient to decide about the semantic scope of interrogatives expressing QUANTITY (MASS–COUNT). Nevertheless, some clear examples can be given in order to illustrate possible patterns (Table 6.10).

Some languages have only one (e.g., Kolyma Yukaghir), others have two different forms (e.g., English, Mandarin, Mongolian). If there are two different forms, these may either be related etymologically (e.g., English, Mongolian) or can have a completely different origin (e.g., Mandarin). In some cases COUNT is derived from MASS (e.g., German), which appears to be the unmarked category. In some other cases the semantic scopes of individual forms overlap (e.g., Mandarin). The use of any of the forms is usually based on subtle differences and the boundary between mass and count nouns is language-specific. In principle, the distribution of different types, such as MASS=COUNT vs. MASS≠COUNT, or SELECTION=PLACE vs. SELECTION≠PLACE could be shown on geographical maps, but the information for most languages was simply insufficient.

Table 6.9: Some examples for the semantic scope in the category PLACE. Only a selection of forms and languages is listed

Language	LOCATION	DIRECTION	SOURCE
Amdo Tibetan	kaŋ-na	kaŋ-a	kaŋ-ni
Buryat	xaa-(na)	xai-sha	xaana-haa
Chukchi	miŋ-ke	miŋ-kəri	meŋ-qo(rə)
CS Yupik	na-ni	na-vek	na-ken
English	where	where (to)	where from
Evenki	ii-du	ii-le	ii-duk
German	wo	wo-hin	wo-her
Japanese	doko (ni)	doko e/ni	doko kara
Ket	bisəŋ	biltàn	bifil
Khakas	xay-da	xay-ya	xay-daŋ
Kolyemal	ɔdi-mæ	ɔdi-ř	ɔdi-sɔ
Kolyma Yukaghir	qo-n	qa-ŋide	qo-t
Manchu	ai-ba-de	ai-ba-de	ai-ba-ci
Mandarin	(zài) nǎ-li	(wǎng) nǎ-li	cóng nǎ-li
Tundra Nenets	xə-n'a-na	xə-n'a-h	xə-n'a-d°
Ukrainian	de	kudý	zvidky

Table 6.10: Some examples for the semantic scope in the category QUANTITY

Language	MASS	COUNT
English	how much	how many
German	wie viel	wie viel-e etc.
Kolyma Yukaghir	qamun	qamun
Mandarin	duōshǎo	duōshǎo, jǐ- + CLF
Mongolian	xedii	xedn

### 6.2.3 Diachrony of interrogatives

§4.3 identified seven possible diachronic developments of interrogatives. Of these, the convergence of forms is apparently only attested in the northern Tungusic languages Oroqen and Khamnigan Evenki, where the two Proto-Tungusic interrogatives \**ja* ‘which, what’ \**Kai* ‘what’ coalesced in a form *i(i)*- (§5.10.3). The replacement of interrogatives as in Italian *che* > *che cosa* > *cosa* ‘what’ or the development of interrogatives from scratch do not appear to be very widespread. However, some examples can perhaps be found in Tocharian, e.g. PIE \**k<sup>w</sup>i-* ‘INT’ + \**so* ‘DEM’ > Proto-Tocharian \**k<sup>w</sup>asē* > *k<sub>u</sub>se* > *se* ‘who’ (§5.5.3.5). The remaining four developments (repeated here in Table 6.11) are more frequent.

Table 6.11: The most important diachronic developments of interrogatives

	Schematic	Details	Example
1	INT <sub>1</sub> > INT <sub>1</sub>	phonological changes	PIE * <i>kwód</i> > OE <i>hwæt</i> > NE <i>what</i>
2	INT <sub>1</sub> > INT <sub>2</sub>	semantic changes	Wutun <i>age</i> ‘which (one) > who’
3	INT <sub>1</sub> -X <sub>GRAM</sub> > INT <sub>2</sub>	inflection (> fusion)	English <i>where to</i>
4	INT <sub>1</sub> -X <sub>LEX</sub> > INT <sub>2</sub>	derivation, reinforcement (> fusion)	English <i>how much</i>

(1) Most languages have a large number of inherited interrogatives. An exception is Mandarin, which apparently preserves only the two Old Chinese interrogatives *shéi* (*shuí*) 谁 and *jǐ*- 几 (§5.9.3.1). For details of individual language families, the reader is referred to Chapter 5. The loss of the resonance due to phonological changes, which may lead to a very different interrogative systems, is only attested for Tungusic languages (§5.10.3).

(2) Semantic changes appear to be quite infrequent, but they are often difficult to detect because of the lack of data. Some relatively clear examples have been collected in Table 6.12, which includes only those cases that do not also involve inflection or derivation. For example, Mandarin *gàn shénme* can mean both ‘to do what’ and ‘why’, without requiring any additional marking.

There are too many instances of inflection (3) or derivation (4), which is why Table 6.13 lists only general patterns illustrated with some examples. Over time many derived interrogatives fuse, are subject to phonetic erosion, and become unanalyzable (e.g., MHG *wār + umbe* ‘where + around’ > German *warum* ‘why’).

Table 6.12: Changes in the semantic scope of interrogatives with some examples

Source	Target	Examples
ACTIVITY	→ REASON	Mandarin <i>gàn shénme</i> ‘to do what, why’
MANNER	→ REASON	Mandarin <i>zěnme</i> ‘how, why’
PLACE	→ MANNER	Manchu <i>absi</i> ‘whither, how’
	→ REASON	German <i>woher</i> ‘whence, why’
SELECTION	→ PERSON	Solon <i>awu</i>
THING	→ REASON	?Chitose Ainu <i>hemanta</i> ‘what, why’

Table 6.13: Some possibilities of inflection and derivation of interrogatives

Type	Examples
INT + ADJ	Japanese <i>do-no</i> ‘which’ etc.
INT + ADP	English <i>where to</i> , Mandarin <i>wǎng nǎli</i> ‘id’ etc.
INT + CASE	Evenki <i>ii-le</i> ‘whither’ etc.
INT + CLF	Mandarin <i>nǎ(-yi)-ge</i> , <i>jǐ-ge</i> etc.
INT + CVB	Manchu <i>ai.na-me</i> ‘why’, Khorchin <i>ᠵᠣᠰ ᠭ᠎ᠠ</i> , Mongolian <i>yaa-j</i> ‘how’, <i>yaa-gaad</i> ‘why’ etc.
INT + DEM	Tocharian * <i>k<sup>w</sup>əṣē</i> , Slavic * <i>къто</i> , Lingshi Mandarin <i>ue<sup>44</sup>ṣu<sup>44</sup></i> ‘who’
INT + gender	Ket <i>biseŋ-du</i> ‘where-3SG.M’ etc., German <i>welch-</i> ‘which’, Ukrainian <i>kotoryj</i> ‘which’ etc.
INT + N	English <i>what kind</i> , Mandarin <i>zěnme yàng(de)</i> ‘id’ etc.
INT + number	Aleut <i>kiin</i> ‘who (SG)’, <i>kiin-kux</i> ‘who (DU)’, Eastern <i>kiin-kun</i> , Atkan <i>kiin-kus</i> ‘who (PL)’ etc.
INT + numeral	English <i>which one</i> , Written Tibetan <i>ci.cig</i> , Mandarin <i>nǎ-yi-</i>
INT + person/number	Ket <i>biseŋ-di</i> ‘where-1SG’ etc., Nivkh <i>ja(ŋ).gu-</i> ‘how’, German <i>welch-</i> ‘which’, Ukrainian <i>kotoryj</i> ‘which’ etc.
INT + Q	Korean <i>nwu-ku</i> ‘who’, Tocharian B <i>tāssi &lt; tā + aṣsi</i> ‘where’, Amur Nivkh <i>aŋ ? &lt; nar-ŋa</i> ‘who’
INT + V	English <i>to do what</i> , Manchu <i>ai-na-</i> ‘id’ etc.



6.3 The significance of the grammar of questions

Table 6.14: Possible instances of borrowing of interrogatives in NEA

Source		Target
Mongolic *kedü- ‘how many’, *keli ‘when’, *yaxun ~ *ya- ‘what’	?→	Tungusic *Kadü, *Kaali, *ja(-kun)
Dagur joonde ‘why’	→	Nanmu Oroqen joonde
Mongolic *ya- via Tungusic *ja- ‘(to do) what’	?→	Nivkh ja-
Amur Nivkh řag ‘where’, East Sakhalin Nivkh nunt/nud ‘what’	?→	Uilta saa ‘where’, sado ‘where’, nuulu ‘whither’
Turkic, e.g. Tuvan kay(i) ‘which’, kařan ‘when’	→	Selkup qaj, ?quřan
Solon awu ‘who’	→	Nanmu Oroqen awu
Udegheic, e.g. Udihe ni ‘who’, adi, ‘how’, je’-(u), je-me ‘what kind’, Oroch oni ‘how’, Udihe ono-bui ‘which’	→	Kilen ni, adi, ja, yao, iama-, oni, ? ɔni-biçi
Ewenic, e.g. Evenki ři ‘who’, ady ‘how much’, i-du ‘where’, ee-da ‘why’, Oroqen jEema ‘which’, aali ‘when’	→	Kili řii, adii, i-du, ii-daj, e-ma, ali
Ewenic, e.g. Oroqen oki ‘how much’	→	Kilen uki
Jurchenic, e.g. Manchu ya ‘which’, atanggi ‘when’	→	Kilen iətin
Hezhou Chinese aži <sup>24</sup> gə ~ aji <sup>24</sup> gə ‘which one, who’	→	Mangghuer ayige
Mandarin jidiān ‘what o’clock’	→	Santa dʒidʒiən(-də)
Sinitic, e.g. Mandarin mà ‘what’	→	Santa, Kangjia ma-
Mandarin shénme ‘what’	→	Chinese Pidgin Russian šýma
Nganasan silī ‘who’	→	Taimyr Pidgin Russian syly
Russian kokda ‘when’, kakoj, kotoryj ‘which’, kuda ‘where’	→	Copper Island Aleut ka(g)da, kakuy, katorəye, kuda
Russian kakoj ‘which’	→	Chalkan qaƣoy ~ kakoy
Russian kak ‘how’	→	Selkup kak ~ kař
Iranian, e.g., Sogdian (ə)ču ‘what’	??	Siberian Turkic, e.g. Chalkan t’u(u) etc.
Koreanic *e-	??	Japonic *e-
Turkic *qay-, e.g., Uyghur qay-, Khakas xay- ‘what, which’	??	Tungusic *Kai, e.g., Alchuka kai-, Nanai xai

Consider the following example from a language spoken in Eastern Sulawesi.

- (2) Balantak (Celebic, Austronesian)  
*ime a men mae', yaku' kabai i koo?*  
who ART REL go 1SG or PERS.ART 2SG  
'Who will go, you or I?' (van den Berg & Busenitz 2012: 66)

The context of the utterance is difficult to reconstruct. But the content question contains an interrogative that represents an initial categorization of a given referent (in this case PERSON). The following alternative question represents possible predictions concerning the identity of that referent. The choice of the interrogative thus also offers direct evidence for the most basic categorization and organization of our knowledge.

The Introduction has claimed that the grammar of questions can function as yardstick for measuring the intensity of the intensity of language contact, areal convergence, unusually strong language contact, and simplification. The remainder of this section briefly evaluates these claims and argues that in many cases they give valuable and good results. On the identification of long-range relationships see §6.2.1.

Regarding the Amdo Sprachbund, for example, Slater (2003a: 6) observed the following:

It certainly is true that intense two-language contact situations have resulted in many instances of localized contact-induced language change, and I do not mean to suggest that two-language comparisons should not be made in the Qinghai-Gansu region. However, what has often been lacking is an overview of the regional processes of linguistic feature diffusion.

In fact, as seen in §3.5, many features mentioned by Janhunen (2012c: 180ff.) such as SOV word order fail to define the region as a linguistic area because they are too frequent worldwide and in adjacent regions. However, the investigation of the grammar of questions has potentially revealed two features that could help define the Amdo Sprachbund. Sandman (2012: 384), by comparing two languages, came to the following reasonable conclusion.

In Bonan, the most common interrogative marker is *-u*. The interrogative marker *-mu* is formed by attaching the interrogative marker *-u* to the narrative aspect marker *-m*. The narrative aspect marker indicates stative or habitual aspect in Bonan. The borrowing of the interrogative marker *-mu* is another example of grammatical borrowing from Bonan to Wutun.

By just looking at these two languages the conclusion is, of course, very plausible because the marker has a clear etymology in Bonan but not in Wutun. However, there is another possibility that treats the Wutun question marker as a loan from Turkic, e.g. Uyghur =*mu*, Sarig Yughur =*mu*, or Salar =*mU*, perhaps via Hezhou Chinese =*mu* or Tangwang =*mu*. If this scenario is accurate, the markers in Wutun and Bonan are only

similar by chance. In fact, the question marker *-mu* in Bonan has parallels in other Mongolic languages of the area such as Mongghul *-muu*, Santa *-mu*, and Kangjia *-mü*. However, even if one excludes the Mongolic question marker, the presence of a relatively widespread and specific question marker *=mu* in Turkic and Sinitic languages of the region that is absent in the surrounding area is certainly a better defining feature than SOV word order. Another example is the presence of single marking on the first alternative in alternative questions (Figure 6.10) shared at least by Gangou (not shown on the map), Hezhou, Wutun, Santa, Bonan, Kangjia, and Mangghuer. This feature, again, can also be found in Uyghur as well as Urumqi Hui Chinese, but not in the surrounding languages in NEA. Given the lack of information on alternative questions, this feature might well be more widespread in the area. In fact, there is some indication that it can perhaps also be found in the immediate south of the Amdo Sprachbund (see §4.2.1).

The Amdo area, of course, is known to be a region of strong linguistic convergence and even creolization, but question marking can also identify contact situations that are otherwise hard to detect. It is well-known, for example, that there was contact between Koreanic and the Tungusic language Manchu. However, previous studies have been quite unsuccessful in identifying any conclusive linguistic evidence for this historical fact. Vovin (2013a: 224f.) has collected a short but extremely valuable list of 17 Koreanic items in Manchu, some of which unfortunately are somewhat problematic. In my opinion, the Manchu third person pronoun *i*, for example, more likely derives from Mongolic *\*i* (Janhunen 2003d: 18), because it shares an identical oblique stem formation, e.g. Manchu *in-i* ‘3SG.OBL-GEN’, Proto-Mongolic *\*in-U* > *\*in-i* ‘3SG.OBL-GEN’. Pronouns are not easily borrowed and perhaps only the contact with Mongolic was strong enough (e.g., Doerfer 1985). At least some of his correspondences such as Manchu *fucihi* ‘Bud-dha’ (from Middle Korean *pwüthyë*) are very plausible. In fact, the form *p’ut(‘)ihi.n* in the language Bala makes this even more likely (Mu Yejun 1987), but cultural loanwords such as this are not necessarily a sign of direct language contact. This study has identified a whole list of question markers in Manchu (and Jurchenic) that appear to systematically derive from a Koreanic source. Not only are the Manchu question markers very different in form and semantic scope from the rest of Tungusic, but the forms are strikingly similar to Koreanic (see §5.7.2, §5.10.2). Such markers can only have been borrowed through direct interaction of the speakers of these languages.

Manchu is perhaps the most aberrant Tungusic language and I have previously put forward the possibility that it might even be comparable to languages such as Afrikaans (Hözl 2012; 2015a: 151). In fact, not only the question marking system, but also the interrogative system is rather different from other Tungusic languages (§5.10.3). While Manchu preserves some Tungusic interrogatives, there is a large amount of innovative forms that are based on the two stems *ai* ‘what’ and *ya* ‘which’, which speaks in favor of a certain amount of simplification (Table 6.15) due to massive non-native acquisition in the history of Manchu (McWhorter 2007; Operstein 2015). Of course, this theory should actually include all of Jurchenic.

Another striking example is Mandarin, which is also known to have experienced a certain amount of simplification with respect to Old or Middle Chinese and other Sinitic

languages (McWhorter 2007: 104-137) and contains a large amount of analyzable interrogatives as well (§5.9.3.1). Notice that this is qualitatively different from the contact between Manchu and Koreanic, which lead to complexification instead (Table 6.15). Manchu actually exhibits more question markers than other Tungusic languages and this must be due to influence from Koreanic. Unlike other Tungusic languages, but similar to Korean, Manchu also employs the question markers in content questions, which from a certain perspective could be interpreted as a type of redundancy. This must be the result of a different language contact scenario that involves longstanding contact and perhaps some bi- or multilingualism. See Hölzl (2017c) for an additional discussion of simplification and complexification of Tungusic interrogative systems.

Table 6.15: Complexification and simplification (Trudgill 2011: 62)

Complexification	Simplification
irregularization	regularization of irregularities
increase in opacity	increase in morphological transparency
increase in syntagmatic redundancy	reduction in syntagmatic redundancy
addition of morphosyntactic categories	loss of morphological categories

Tungusic also offers a good example for yet another type of language contact that leads to the mixing of languages. The language Kilen, for example, is well-known to be a mixed Tungusic language and has been sometimes classified with Nanai (e.g., Alonso de la Fuente 2011; Janhunen 2012d; this study) and sometimes with Udihe (Kazama 2003). Influence from Manchu has often been overlooked, however (see Hölzl 2017a). In fact, Kilen exhibits interrogatives that can clearly be shown to derive from Nanai, Udihe, and Manchu, which represent three different branches of the language family. Consider the following example.

- (3) Kilen (Tungusic)  
*ni jarin ja-tulə ənə-kiciə?*  
 who when where-ALL go-?SUBJ  
 ‘Who would like to go to where when?’ (Zhang 2013: 163)

The verb most likely derives from Nanaic, but every interrogative must stem from other branches of the language family (Table 6.16). Such a mixed language can only be the result of multilingualism throughout the entire speech community: “Unlike creoles, mixed languages arise in bilingual settings in which the speakers are equally fluent in the two codes.” (Operstein 2015: 6)

Another example seen in NEA is Copper Island Aleut, which exhibits Russian and Aleut interrogatives (§5.4.3). Creoles are both mixed and exhibit “extreme simplification on all levels” (McWhorter 2007: 254) due to non-native acquisition. In principle, they

6.3 The significance of the grammar of questions

Table 6.16: The etymological background of the Kilen elements in (3)

Language	who	when	whither	ALL	to go
Manchu (Jurchenic)	we	ya erin	ya-de	-de	gene-
Nanai (Nanaic)	ui	xaali	xaosi	-dola	ənə-
Udihe (Udegheic)	ni	ali	j'e-uxi	-dulA	ɲene-
Evenki (Ewenic)	ɲi	ookin	i-le	-dulA/- tulA	ɲene-

should exhibit both a simplified or transparent interrogative system as well as interrogatives from different sources and this indeed seems to be the case for at least some of them (Bickerton 2016 [1981]: 65f.; Muysken & Smith 1990). There are no true creole languages in NEA, but Taimyr Pidgin and Chinese Pidgin Russian had interrogatives of Russian and dialectal Russian origin as well as at least one from Nganasan and Chinese, respectively (§5.5.3.3). Taimyr Pidgin furthermore had at least some innovative interrogatives such as *kudy-mera* 'where', *kudy-mesto* 'where', and *kakoj storona* 'whither'.

Another type of change we see under creolization is the translation of individual forms such as Chinese Pidgin Russian *mnogo-malo*, which consists of Russian *mnógo*/МНОГО 'much', *málo*/МАЛО 'little' and is a direct translation of Chinese *duōshǎo* 多少 'how much'. Calques are not necessarily restricted to creole and pidgin languages, however, but can also be found in instances of bilingual contact. Most cases found in NEA are partial borrowings and contain an autochthonous interrogative, e.g. Qiang *na-tian* from Chinese *jǐ diǎn* 几点 'what hour' (LaPolla & Huang Chenglong 2003: 53f.). A mixture of calque and borrowing can also be found in Santa *yan shihou* from Mandarin *shénme shíhou* 'what time'. Special cases of borrowing are, furthermore, *iamə-dzaka* 'what thing' and perhaps *iama-ərin* 'what time' in the Tungusic language Kilen, which derive from two different sources. The actual interrogative has been borrowed from Udihe *je-me* 'what kind', while the second elements derive from Manchu *jaka* 'thing' and perhaps *erin* 'time', both of which are also present in Manchu interrogatives. In some cases an interrogative has been entirely translated. Manchu *ai se-me* and Khorchin Mongolian *jəv gə-ʃ*, for example, have the same underlying pattern 'what say-CVB.IPFV' and both mean 'why'. Khalkha *xer olon* or Ket *bilon* appear to have been formed on the basis of a European pattern also seen in English *how many/much*.

These examples illustrate that the grammar of questions can indeed function as a preliminary but valuable tool for the identification of different types of language contact, but this section has focused on individual instances of diffusion or convergence, exclusively. The following maps of the atlas allow an additional identification of large patterns of areal convergence that is impossible from the study of individual languages alone.

## 6.4 An atlas of the grammar of questions in Northeast Asia

The geographical extent of certain features will be demonstrated with the help of a synchronic sample of 83 languages (Figure 6.4, Table 6.17) that covers all 14 language families of NEA. Languages with a wide geographical distribution are underlined in Figure 6.4 and shown with bigger symbols in Figure 6.5-Figure 6.16 below. The maps exclude extinct languages and list only some dialects of a given language. An exception is made for Ainuic, which by now is probably completely extinct but has been added for reasons of completeness.

Given somewhat unclear boundaries between languages and dialects in Japonic, dialectal variation in this family may be slightly overrepresented. It may be noted that the lack of data for some languages might have led to some distortions. Nevertheless, the general areal patterns seem to be valid. The white line in Figure 6.4 indicates the rough definition of NEA adopted in this study. The distribution of the languages clearly shows a large spread zone over large parts of Northeast Asia, including Northern China, Mongolia, Siberia, Korea, and Japan (excluding Hokkaidō and the Ryūkyūan Islands) with few but widespread languages. Residual zones with many local languages are found in northern Manchuria (including Sakhalin and Hokkaidō), the Ryūkyūan Islands, the Aleut Islands, Amdo, the Altai, along the Yenisei, and perhaps on Kamchatka.

6.4 An atlas of the grammar of questions in Northeast Asia

Table 6.17: The synchronic sample of 83 languages used for the maps

Family	Language	Number
<b>Ainuic</b>	Chitose Ainu	1
	Saru Ainu	2
	Shizunai Ainu	3
<b>Amuric</b>	Amur Nivkh	4
	East Sakhalin Nivkh	5
<b>Chukotko-Kanchatkan</b>	Chukchi	6
	Alutor	7
	Koryak	8
<b>Eskaleut</b>	Itelmen	9
	Central Siberian Yupik	10
	Aleut, Atkan	11
<b>Indo-European</b>	Copper Island Aleut	12
	Russian	13
	Ukrainian	14
<b>Japonic</b>	Plautdiitsch	15
	Yiddish	16
	Sarikoli	17
	Japanese	18
	Hachijō	19
	Yilan Creole	20
	Yuwan	21
	Okinoerabu, Masana	22
	Shuri	23
	Ikema	24
Ōgami	25	
<b>Koreanic</b>	Irabu	26
	Hateruma	27
	Hatoma	28
	Miyara	29
	Sonai	30
<b>Mongolic</b>	Korean	31
	Jeju	32
	Dagur	33
	Buryat	34
	Khamnigan Mongol	35
	Cyrillic Khalkha	36
	Oirat	37
	Shira Yughur	38
	Santa	39
	Bonan	40
	Kangjia	41
	Huzhu Mongghul	42
	Minhe Mangghuer	43

6 *Interrogative constructions in Northeast Asia: A summary*

---

<b>Trans-Himalayan</b>	Standard Mandarin	44
	Urumqi Hui Mandarin	45
	Hezhou	46
	Tangwang	47
	Wutun	48
	Amdo Tibetan, Gonghe	49
	Baima	50
	Zhongu	51
<b>Tungusic</b>	Even	52
	Evenki	53
	Negidal	54
	Kh. Evenki	55
	Oroqen, Xunke	56
	Solon, Huihe	57
	Udihe	58
	Kilen (An Jun)	59
	Nanai	60
	Ulcha	61
	Uilta	62
<b>Turkic</b>	Sibe	63
	Salar	64
	Kazakh	65
	Kyrgyz	66
	Tatar (Chinese)	67
	Uyghur	68
	Tuvan	69
	Tofa	70
	Khakas	71
	Sarig Yughur	72
	Altai Turkic	73
	Chulym, Middle	74
	Dolgan	75
Yakut	76	
<b>Uralic</b>	Nganasan	77
	Forest Enets	78
	Tundra Nenets	79
	Selkup, Taz	80
<b>Yeniseic</b>	Ket	81
<b>Yukaghiric</b>	Kolyma Yukaghir	82
	Tundra Yukaghir	83

---

6.4 An atlas of the grammar of questions in Northeast Asia

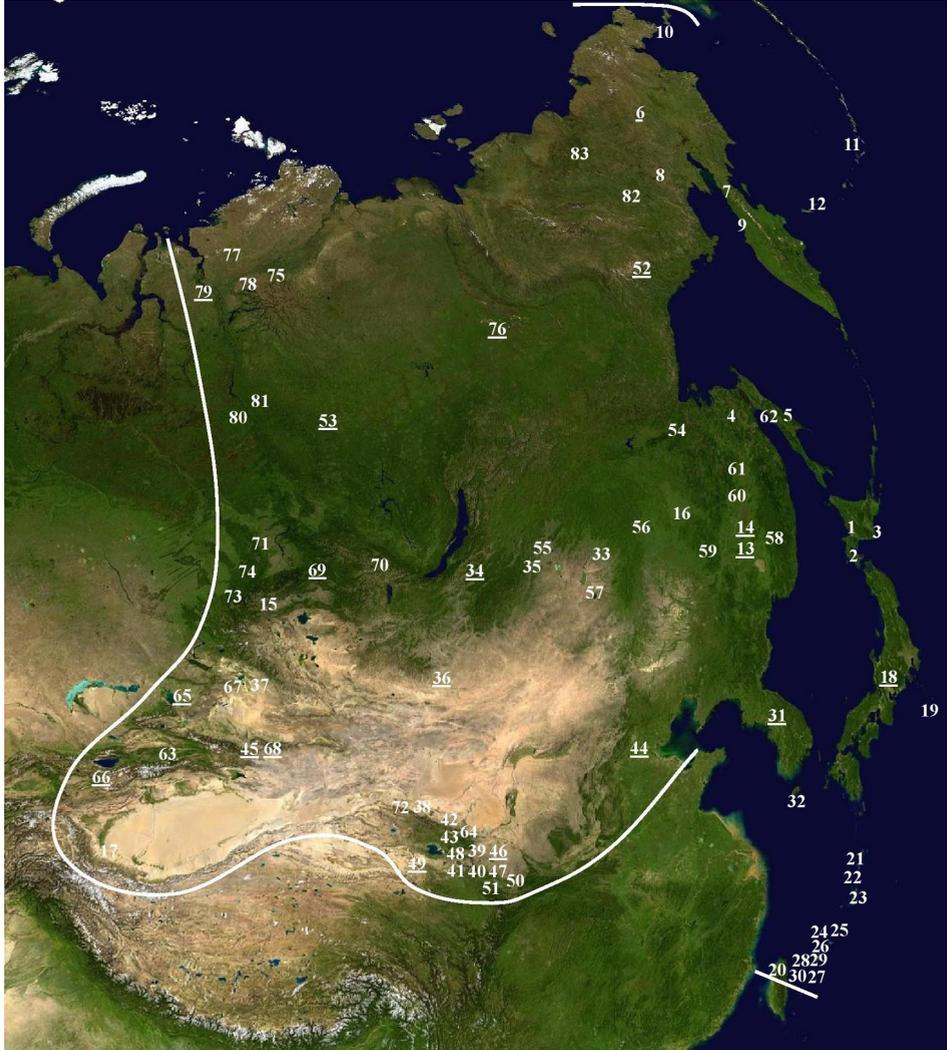


Figure 6.4: Approximate geographical location of the 83 languages in the sample (1)

6 Interrogative constructions in Northeast Asia: A summary



Figure 6.5: Polar question marking (2)

Color	Type	Languages
white	sentence-final marker	45
white with black dots	mixed	19
black with white outline	suffix(es)	4
gray with white outline	no morphosyntactic marker	6
white with vertical black lines	sentence initial marker	3
white with horizontal black lines	prefix(es)	2
white with diagonal black lines	second position marker	2
black	preverbal marker	1
light gray with black outline	word order	1
Total	9	83



Figure 6.6: Sentence-final polar question marker present (3)

Color	Type	Languages
white	sentence-final marker	62
gray with white outline	not present	21
Total	2	83

6 Interrogative constructions in Northeast Asia: A summary

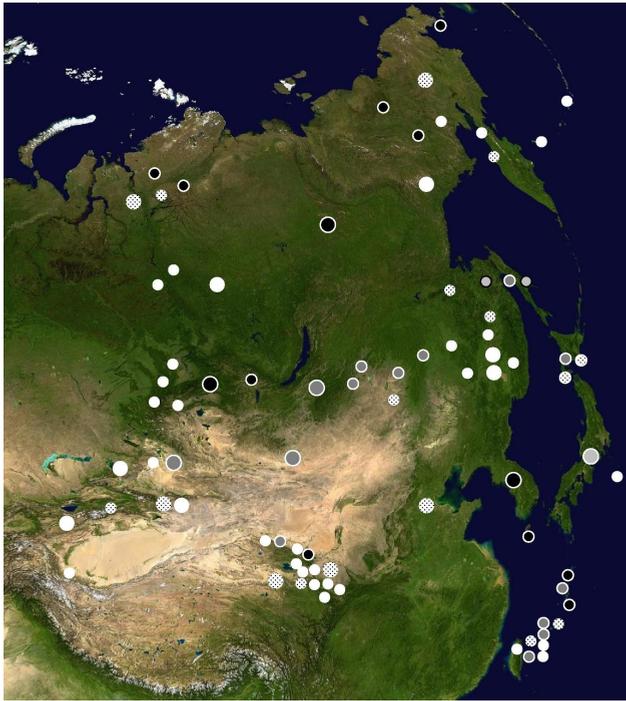


Figure 6.7: Content question marking (4)

Color	Type	Languages
white	no morphosyntactic marker or unclear	36
gray with white outline	sentence-final marker	15
black with white outline	suffix(es)	13
white with black dots	mixed	14
light gray with black outline	mobile enclitic	2
Total	5	83

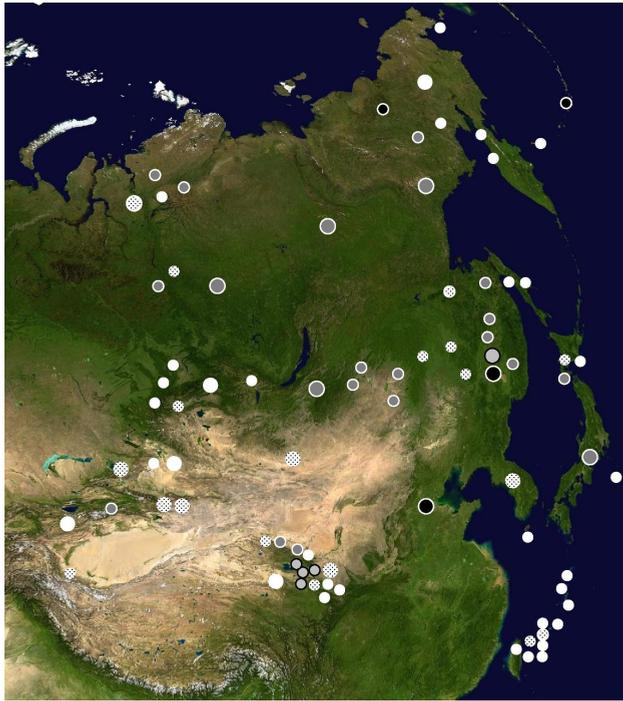


Figure 6.8: Alternative question marking (5)

Color	Type	Languages
white	unclear	34
gray with white outline	double marking	21
white with black dots	mixed, other	19
light gray with black outline	single marking	5
black with white outline	disjunction	4
Total	5	83

6 *Interrogative constructions in Northeast Asia: A summary*

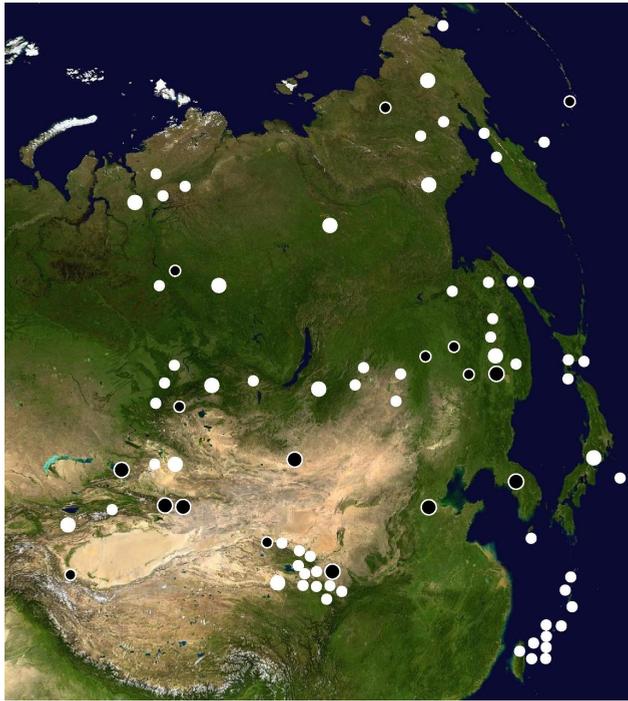


Figure 6.9: Presence of disjunction in alternative questions (6)

Color	Type	Languages
white	unclear or not present	65
black with white outline	present	17
Total	2	83

6.4 An atlas of the grammar of questions in Northeast Asia

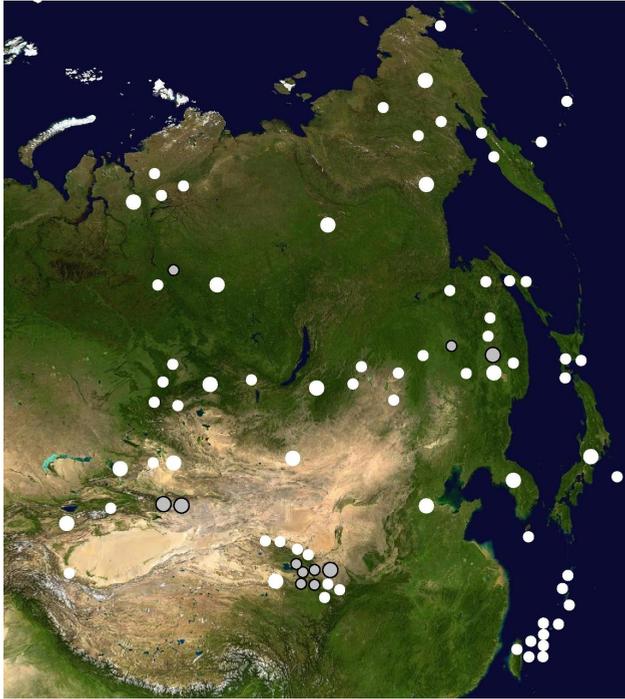


Figure 6.10: Presence of single marking in alternative questions (7)

Color	Type	Languages
white	unclear or not present	72
light gray	present	11
Total	2	83

6 Interrogative constructions in Northeast Asia: A summary

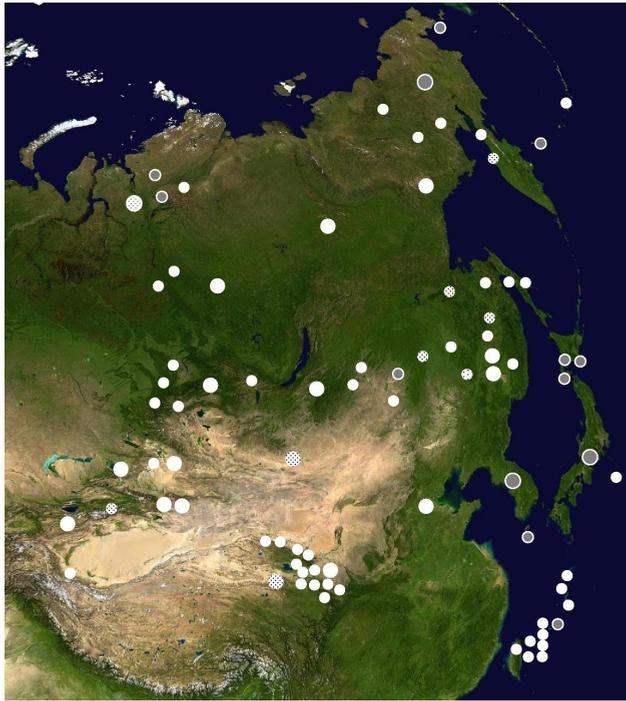


Figure 6.11: Polar versus content question marking (8)

Color	Type	Languages
white	different	61
gray with white outline	identical	13
white with black dots	mixed	9
Total	3	83

6.4 An atlas of the grammar of questions in Northeast Asia



Figure 6.12: Polar and content questions overtly marked differently (9)

Color	Type	Languages
white	no or unclear	62
gray with white outline	yes	21
Total	2	83

6 Interrogative constructions in Northeast Asia: A summary

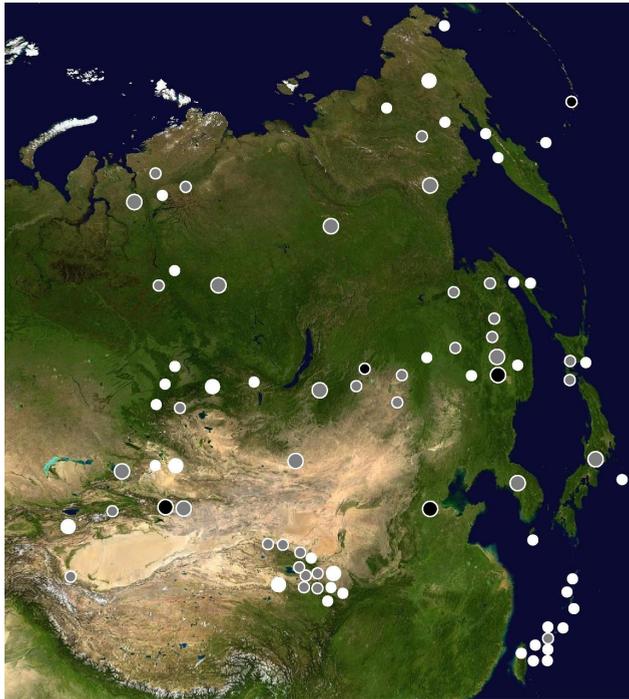


Figure 6.13: Polar versus alternative question marking (10)

Color	Type	Languages
gray with white outline	identical	37
black with white outline	different	5
white	mixed or unclear	41
Total	3	83

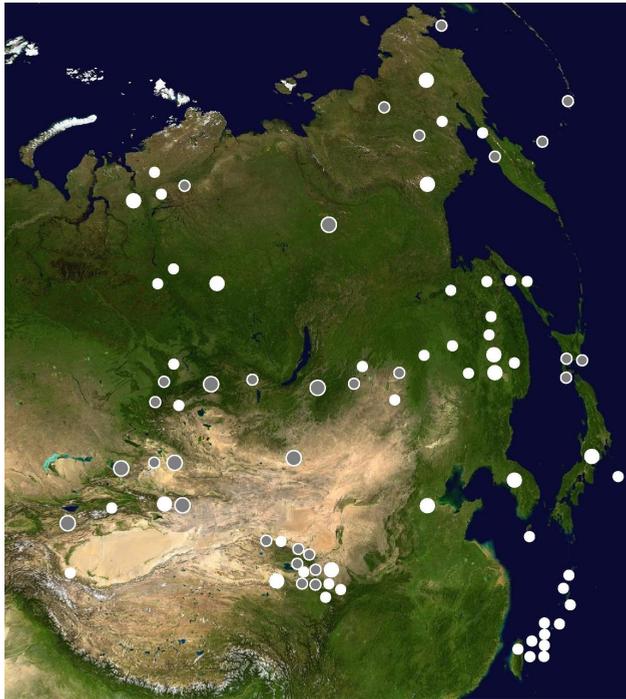


Figure 6.14: KIN-interrogatives (11): The interrogative meaning ‘who’ in a given language has the form KIN (velar or uvular plosive or fricative, (high) vowel (short or long), (apical) nasal), followed by an optional final vowel, e.g. Turkish *kim*, Forest Nenets *kim’a*, Aleut *kiin* etc.

Color	Type	Languages
white	not present	53
gray with white outline	present	30
Total	2	83

6 Interrogative constructions in Northeast Asia: A summary

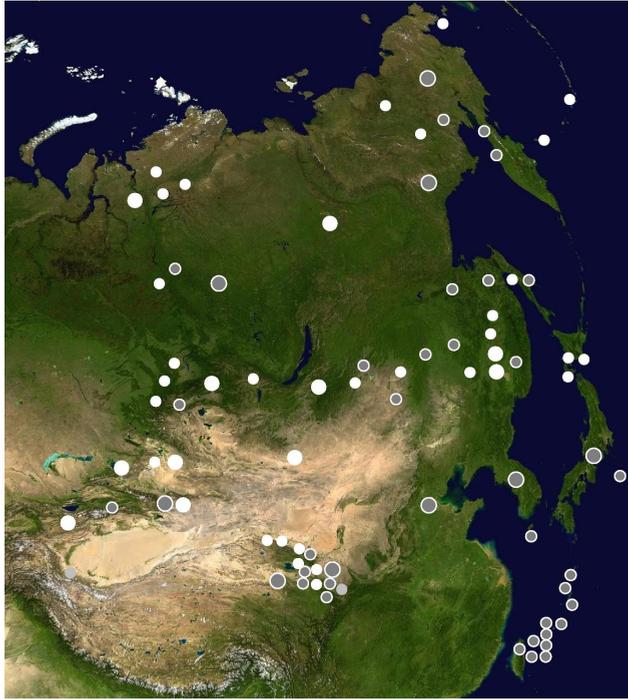


Figure 6.15: K-interrogatives (12): More than two interrogatives in a given language start with the same velar or uvular plosive or fricative, e.g. Nanai *xai* 'what', *xado* 'how many', *xooni* 'how', Uyghur *qaysi* 'which', *qačan* 'when', *qandaq* 'how' etc.

Color	Type	Languages
gray with white outline	not present	42
white	present	39
light gray	unclear	2
Total	3	83

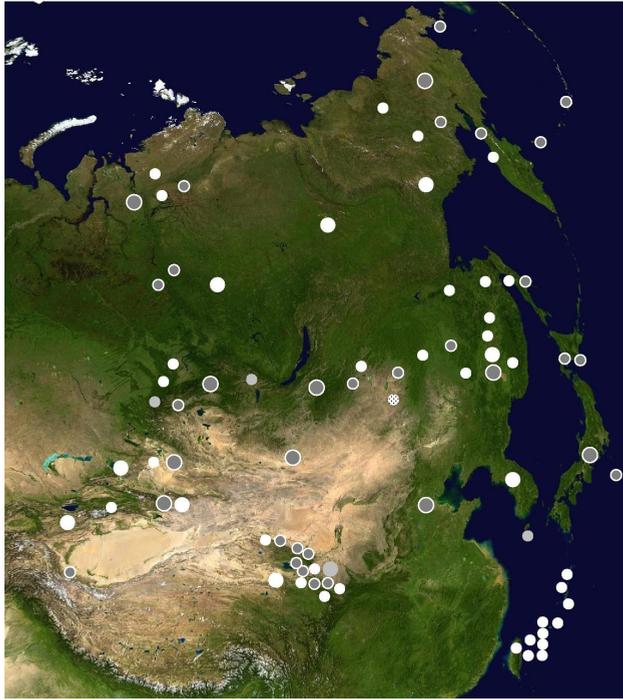


Figure 6.16: The personal interrogative ‘who’ has a different initial consonant from all other interrogatives (13)

Color	Type	Languages
white	yes	43
gray with white outline	no	35
light gray	unclear	4
white with black dots	mixed	1
Total	4	83



## 7 Conclusion

According to Evans & Levinson (2009: 429), “we are the only species with a communication system that is fundamentally variable at all levels.” The investigation of linguistic diversity thus should be a major concern of linguistics in general and of typology in particular. The main research question of this study was, following Bickel (2007: 248), “what’s where why?”

Asking “what’s where?” targets universal preferences as much as geographical or genealogical skewings, and results in probabilistic theories stated over properly sampled distributions. Asking “why?” is based on the premises that (i) typological distributions are historically grown and (ii) that they are interrelated with other distributions. (Bickel 2007: 239)

Therefore, the present study is not a classical synchronic typological investigation, but focused on the distribution and preliminary explanation of linguistic diversity found in the limited geographical area of Northeast Asia (NEA), tentatively defined as the area north of the Yellow River and east of the Yenisei (e.g., Chard 1974). Another question formulated in the Introduction was whether the concept of Northeast Asia makes sense from the point of view of areal linguistics. The answer is certainly yes, but with limitations. The definition of Northeast Asia as a concept strongly depends on its opposition with Mainland Southeast Asia (MSEA, Enfield & Comrie 2015). Regarding the number of languages (language diversity), NEA with perhaps up to 150 languages ranks much lower than MSEA, which is the home of up to 600 different languages. In terms of different linguistic stocks, however, NEA has 14 instead of only 5 found in MSEA (phylogenetic diversity). In comparison, the region of New Guinea is home to approximately 1200 languages from about 35 language families on an area of only 850,000 km<sup>2</sup> (Foley 2000). In NEA, Mongolia alone is larger than that area. There are similarly pronounced regional differences in linguistic diversity within Northeast Asia. The highest concentration of languages can be found in peripheral regions such as the Amdo region, the Ryūkyūan Islands, in the Amur river shed, and around the Altai extending northwards along the Yenisei as well as southward along adjacent mountainous regions. Following Nichols (1992; 1997) these can be characterized as *residual* or *accretion zones*. Language diversity is at its lowest in central parts around Mongolia, northern China, central Siberia, Korea, and central parts of Japan, which qualifies as a large coherent *spread zone*. Regarding phylogenetic diversity, there is quite a different distribution that peaks around the eastern part of NEA along the Pacific Rim (Pacific NEA), where representatives of 12 of the 14 language families of NEA can be found (e.g., Anderson 2010). Historically, however, both Yeniseic and Samoyedic, which are the only exceptions, have been spo-

## 7 Conclusion

ken further towards the southeast as well. No doubt there is a multitude of reasons for these strong differences in linguistic diversity, including climatic (e.g., temperature, precipitation), geographical (e.g., landscape roughness, river density), and cultural factors (e.g., subsistence patterns, agriculture, pastoralism, hunting and gathering) (e.g., Nichols 1992; Nettle 1999; Axelsen & Manrubia 2014). Not only is there a complex mixture of different causes located on different time scales, but the importance of individual factors varies from region to region. These factors clearly also influence the size of languages, which is greatest in the southeast (Mandarin, Japanese, Korean) and decreases towards the west and especially towards the north and seems to be directly correlated with the distribution of population density and environmental factors such as climate and vegetation, and, consequently, the existence of agriculture. Understandably, the exact causes of phylogenetic and language diversity could not be investigated within this study, which focused on structural diversity, more precisely the diversity found in the *grammar of questions*, i.e. those aspects of any given language that are specialized for asking questions. The primary distinction made in the grammar of questions of a given language is in question marking, interrogatives, and optional additional functional domains such as coordination, focus, and negation. A comparison of the structural diversity of the grammar of questions found in MSEA and NEA was not feasible as there are simply too many languages to investigate in MSEA. The obvious next step should thus be to expand the typology proposed in this study to Mainland Southeast Asia (see Clark 1985; Huang 1996; Huang et al. 1999; Enfield 2010; Rajasingh 2014, etc.) and to other regions from around the globe. Nevertheless, there is evidence that NEA and MSEA, despite manifold differences (Chapter 3), together form one large area with a preponderance of sentence-final polar question markers (Dryer 2013). For reasons of space this study necessarily also excluded responses and answers, which is yet another avenue for further research (e.g., Enfield et al. 2010). Future studies should also pay more attention to intonation in questions, which was for the most part neglected here for mere lack of information (e.g., Sicoli et al. 2014 and references therein). However, this study identified many important aspects of the grammar of questions in NEA and beyond, ranging from general principles (Chapter 4) to specific aspects of the languages of NEA (Chapters 5, 6). Given the focus on one area, the typology of questions presented in this study was necessarily limited. I intend to elaborate on it in future studies with a global coverage. For example, the exact distribution and explanation of KIN- and K-interrogatives can only be settled with the help of a global sample of languages. The total discussion mentions over 450 languages and dialects from NEA and beyond (see the Language Index). Altogether about 900 glossed examples were given. The aim was to achieve both a cross-linguistically plausible typology and a resolution of the linguistic diversity of Northeast Asia as much as possible (Voegelin & Voegelin 1964: 2).

Chapters 3 and 6 identified several important areal features such as KIN- and K-interrogatives that have a strong basis in Northeast Asia as well as more localized instances of diffusion and convergence such as in the so-called Amdo Sprachbund. Concerning the grammar of questions, the Tungusic languages play a less important role for NEA than was assumed in §3.4. However, there is no point in arguing whether Northeast

Asia qualifies as a clear-cut *linguistic area*, given the problematic status of the concept itself (e.g., Campbell 2006). In terms of structural diversity, Northeast Asia admittedly has a relatively clear boundary towards the southeast, i.e. Mainland Southeast Asia (e.g., Enfield & Comrie 2015), but not towards the west (e.g., Heggarty & Renfrew 2014a). While there are certain features such as the existence of front rounded vowels that are relatively widespread in NEA, these can often also be found in the adjacent regions towards the west, such as Europe. The reason for this seems to be in the fact that NEA over millennia was the starting point for a multitude of population movements and linguistic spreads over all of northern Eurasia towards the west (e.g., Nichols 1997: 376f.). Another major direction of spread was from southern NEA towards the north, often following the rivers Yenisei and Lena (e.g., Skribnik 2004: 151). Not only do all three large language families of Europe, Indo-European, Uralic, and Turkic, derive from a location further to the east or even from NEA, but the ancestors of *all* native Americans and their languages necessarily had their origin within NEA as well (e.g., Llamas et al. 2016 and references therein). Northeast Asia thus holds a key position for regions as far apart as western Europe and the Americas. One of the best examples for the importance of especially southern NEA for the dispersal of peoples and languages is the recent discovery of the so-called Mal'ta specimen found near lake Baikal that is about 24,000 years old (Raghavan, Skoglund, et al. 2014). It represents a population called the *Ancient North Eurasians* that lack a closer relation to modern East Asians. Instead, Ancient North Eurasians are one of four major founding lineages thus far identified for modern Europeans in the west (Jones et al. 2015), and also significantly contributed to the genome of the Kets along the middle Yenisei in the north (Flegontov et al. 2016) as well as of native Americans that initially spread towards Beringia in the northeast (Raghavan, Skoglund, et al. 2014). Even though the time scales involved are too large to be accessible through historical linguistics, such population movements certainly were also connected with the spread of languages. Take the Yamnaya culture in the Pontic-Caspian steppe, for example, which is thought to have brought both the ANE genome as well as the Indo-European languages into Europe (Anthony 2007; Anthony & Ringe 2015; Allentoft et al. 2015; Haak et al. 2015; Jones et al. 2015). While NEA played a crucial role in the spread of populations to other parts of the world, it was itself reached by populations and thus most likely by languages from as far south as southern China (Hong Shi et al. 2013) and Southeast Asia or perhaps Australasia (Raghavan et al. 2015; Skoglund et al. 2015; Reich 2018: 176-181), which again left traces as far apart as northern and eastern Europe and South America, respectively. These results have potential implications for the search of long-term relations between languages that cannot be restricted to NEA alone.

The title of this study promised *an ecological perspective* and the Introduction tentatively identified the approach as a so-called *ecological typology*. This approach shares its appreciation of human and linguistic diversity with several other approaches (e.g., Evans & Levinson 2009; Levinson 2012b), but in addition follows the so-called *ecological commitment* (Hözl 2015d: 186) that the description of language “should be reconcilable with what is known from ecological research”, which was formulated in analogy to the well-known *cognitive commitment* that continues to define Cognitive Linguistics (e.g., Evans

## 7 Conclusion

2012). While the cognitive approach sees “language as an integral part of cognition” (Langacker 2008: 539), the ecological approach—and what was tentatively called *ecological typology* is only a part of it—in my interpretation sees language as an integral part of ecology, i.e. the *organism-environment system* (e.g., Järvillehto 1998; Odling-Smee & Laland 2009). A similarity of the two approaches is the attempt to find explanations in general principles (Hölzl 2015b: 185), cf. the *generalization commitment* in Cognitive Linguistics (e.g., Evans 2012). In my eyes, *ecology* is a valuable cover term for an emerging field of investigations that, for the explanation of linguistic diversity and language structure, acknowledges a multitude of different *reasons* (e.g., Steffensen & Fill 2014; Bickel 2015; De Busser 2015) that take effect on different *time scales* or *causal frames* (e.g., Enfield 2014). This conceptual shift promises deep implications of which not even the surface could be scratched by this study. Linguistic diversity cannot be considered independently of a multitude of factors, ranging from the invention of the wheel, over the domestication of the reindeer or the biochemistry of the brain, up to the amount of precipitation.

In one sense that was emphasized throughout this book, ecology “represents a shift of emphasis from a single language in isolation to many languages in contact.” (Voegelin & Voegelin 1964: 2) Following Steffensen & Fill (2014), this was called *symbolic ecology*. The subheading *An ecological perspective* thus mainly refers to the aspect of language contact within the entire linguistic landscape of Northeast Asia. The influence of other ecologies such as those mentioned in the Introduction (e.g., cognitive, natural, sociocultural) are only beginning to be understood and consequently had a subordinate position (e.g., De Busser 2015). Nevertheless, there are indications that these influences should not be underestimated and deserve further research (e.g., Axelsen & Manrubia 2014; Everett et al. 2016). An investigation of the impact of climate, for instance, is necessarily based on a global sample of languages which could not be achieved within this regional study. However, a comparison of the results for NEA in this study and a global sample by Dryer (2013j) suggests a possible climatic influence on question marking and especially intonation: The lack of languages in NEA that mark polar questions with intonation alone and do not have further question marking strategies (but see §5.3.2) could be attributed to the fact that, for some reason, such languages are usually located in the tropics. In fact, Everett et al. (2015: 1322) recently found more convincing evidence for a possible climatic influence on language structure:

The sound systems of human languages are not generally thought to be ecologically adaptive. We offer the most extensive evidence to date that such systems are in fact adaptive and can be influenced, at least in some respects, by climatic factors. Based on a survey of laryngology data demonstrating the deleterious effects of aridity on vocal cord movement, we predict that complex tone patterns should be relatively unlikely to evolve in arid [and cold] climates.

In many cases such as this there may be several reasons for a certain phenomenon. Concerning the occurrence of tones in MSEA but not in NEA there are further possible explanations, including language contact or even the occurrence of certain genes (Dediu 2011). Of course, a language can only mark questions with the help of tones if the language possesses tones in the first place (Hyman & Leben 2000: 593).

Additionally, §4.4 has tentatively proposed an ecological theory of questions, which describes them as a form of *exploratory behavior* (e.g., Gibson 1988) in the *dialogical array* (Gibson 1979; Hodges 2009). This exploration can be explained with *specific epistemic curiosity* (Berlyne 1954; Loewenstein 1994), which itself is evoked by so-called “collative” (i.e. novel, changing, complex, conflicting, surprising, or uncertain, Berlyne 1978) properties of the organism-environment system (Järvillehto 1998; Turvey 2009). Humans seek comprehension and clarity, and there are several ways of achieving this, including mental problem solving, physical exploration, or asking questions. However, like other types of exploratory behavior, questions are a proactive process (e.g., Gibson 1988: 5f.). Questions are not merely a request for information, but crucially involve predictions by the speaker and thus depend on our previous experience.



# Appendix A: Data for geographical maps

Numbers refer to maps in the atlas (§6.4). Abbreviations: 2PE = second position enclitic, COP = special copula, D = double marking, dif = different, disj = disjunction, FRV = front rounded vowel, H = high, id = identical, int =intonation, M = mid, ME = mobile enclitic, S = single marking, SFM = sentence final marker, SIP sentence initial particle.

*A Data for geographical maps*

	Language	PQ (2, 3)	CQ (4)	AQ (5, 6, 7)	PQ/CQ (8, 9)	PQ vs. AQ <sup>1</sup> (10)
1	Chitose Ainu	SFM	SFM	D (+D)	id	?id
2	Saru Ainu	SFM, COP	SFM, COP	D	id	?id
3	Shizunai Ainu	SFM, COP	SFM, COP	?	?id	?
4	Amur Nivkh	SFM	ME	D	dif	?id
5	East Sakhalin Nivkh	SFM	ME	?	dif	?
6	Chukchi	-	-, ?prefix	?	?id	?
7	Alutor	SIP	-	?	dif	?
8	Koryak	SIP	-	?	dif	?
9	Itelmen	-	-, ?2PE	?	mixed	?
10	CSY	suffix	suffix	?	id	?
11	Aleut, Atkan	SFM	-	disj	dif	dif
12	CIA	-	-	?	id	?
13	Russian	2PE	-	disj	dif	dif
14	Ukrainian	SIP	-	S	dif	id
15	Plautdiitsch	#V	-	S + disj	dif	id
16	Yiddish	#V + SIP	-	S + disj	dif	id
17	Sarikoli	SFM	-	D + ?disj	dif	id
18	Japanese	SFM	SFM	D	id	id
19	Hachijō	?SFM	?-	?	dif	?
20	Yilan Creole	SFM	-	?	dif	?
21	Yuwan	SFM, suffix	suffix	?	dif	?
22	Okinoerabu, Masana	SFM + suffix	SFM	?	dif	?
23	Shuri	SFM, suffix	suffix	?	dif	?
24	Ikema	SFM	SFM	?	dif	?
25	Ōgami	SFM, suffix	SFM, suffix	?	id	?
26	Irabu	SFM	SFM	D + FOC	dif	?id
27	Hateruma	SFM	-	?	dif	?
28	Hatoma	-	-, SFM	-, D	dif	?mixed
29	Miyara	-	unclear (lack of IND -N)	?	dif	?
30	Sonai	SFM	SFM	?	dif	?

	Language	PQ (2, 3)	CQ (4)	AQ (5, 6, 7)	PQ/CQ (8, 9)	PQ vs. AQ (10)
31	Korean	suffix	suffix	D + disj	id	id
32	Jeju	suffix	suffix	?	id	?
33	Dagur	SFM	?SFM	?D	id	id
34	Buryat	SFM	SFM	D	dif	id
35	Khamnigan Mongol	SFM	SFM	D	dif	id
36	Cyrillic Khalkha	SFM	SFM	D + disj	mixed	id
37	Oirat	SFM	SFM	?	dif	?
38	Shira Yughur	SFM	SFM	D	dif	id
39	Santa	SFM, suffix	-	S	dif	id
40	Bonan	SFM, suffix	-	S/D	dif	id
41	Kangjia	SFM, suffix	-, SFM	S	dif	id
42	Huzhu Mongghul	SFM, suffix	-	D	dif	id
43	Minhe Mangghuer	SFM, suffix	-	S	dif	id
44	Standard Mandarin	SFM	-, (SFM)	disj	dif	dif
45	Urumqi Hui Mandarin	SFM	-, (SFM)	S (+ disj)	dif	dif
46	Hezhou	SFM	-, SFM	S/D (+ disj)	dif	mixed
47	Tangwang	SFM	-	?	dif	?
48	Wutun	SFM	-	S	dif	id
49	Amdo Tibetan, Gonghe	SFM, prefix	-, SFM	?	mixed	?
50	Baima	prefix	-	?	dif	?
51	Zhongu	prefix	-	?	dif	?
52	Even	SFM	-	D	dif	id
53	Evenki	SFM	-	D	dif	id
54	Negidal	SFM + suffix	-, suffix	D (+D)	mixed	id
55	Kh. Evenki	SFM	SFM	D	dif	?dif
56	Oroqen, Xunke	SFM	SFM	?D + disj	mixed	mixed
57	Solon, Huihe	SFM	-, SFM	D	dif	id
58	Udihe	SFM	-	D	dif	mixed
59	Kilen	SFM	-	D, S + disj	dif	mixed
60	Nanai	SFM	-	D	dif	id

*A Data for geographical maps*

	Language	PQ (2, 3)	CQ (4)	AQ (5, 6, 7)	PQ/CQ (8, 9)	PQ vs. AQ (10)
61	Ulcha	SFM	-, ?SFM	D	mixed	id
62	Uilta	SFM	SFM	?	dif	?
63	Sibe	SFM	-, ?SFM	D	mixed	id
64	Salar	SFM	suffix	?	dif	?
65	Kazakh	SFM, suffix	-	D (+ disj)	dif	id
66	Kyrgyz	SFM	-	?	dif	?
67	Tatar (Chinese)	SFM	-	?	dif	?
68	Uyghur	SFM, suffix	-	D/S (+ disj)	dif	id
69	Tuvan	SFM	suffix	?	dif	?
70	Tofa	SFM	suffix	?	dif	?
71	Khakas	SFM	-	?	dif	?
72	Sarig	SFM, suffix	-	D (+ disj)	dif	id
73	Altai	SFM	-	?	dif	?
74	Chulym, Middle	SFM	-	?	dif	?
75	Dolgan	SFM	suffix	D	dif	id
76	Yakut	SFM	suffix	D	dif	id
77	Nganasan	suffix	suffix	D	id	id
78	Forest Enets	-, suffix	-, suffix	?	id	?
79	Tundra Nenets	SFM, -, suffix	-, suffix	D, -	mixed	id
80	Selkup, Taz	preverbal	-	D	dif	id
81	Ket	2PE	-	S/D, disj	dif	mixed
82	Kolyma	SFM	suffix	D	dif	id
83	Yukaghir Tundra Yukaghir	-	suffix	disj	dif	mixed

<sup>1</sup>Question marking here excludes disjunctions and the difference between single and double marking is ignored.

	Language	KIN (11)	K~ (12)	who <sup>2</sup> (13)	FRV	PQ vs. Decl
1	Chitose Ainu	+	+	-	none	4
2	Saru Ainu	+	+	-	none	4
3	Shizunai Ainu	+	+	-	?none	4
4	Amur Nivkh	-	-	+	none	4
5	East Sakhalin Nivkh	-	-	-	none	4
6	Chukchi	-	-	-	none	4 (int)
7	Alutor	-	-	-	none	4
8	Koryak	-	-	-	?	4
9	Itelmen	+	-	+	none	4 (int)
10	CSY	+	+	-	none	3
11	Aleut, Atkan	+	+	+	none	4
12	CIA	+	+	-	?	?4 (int)
13	Russian	-	+	-	none	4
14	Ukrainian	-	+	+	none	4
15	Plautdiitsch	-	-	-	?	4
16	Yiddish	-	-	-	none	4
17	Sarikoli	-	?	-	none	4
18	Japanese	-	-	-	none	4
19	Hachijō	-	-	-	?	?4
20	Yilan Creole	-	-	+	none	4
21	Yuwan	-	-	+	none	4
22	Okinoerabu, Masana	-	-	+	none	4
23	Shuri	-	-	+	none	4
24	Ikema	-	-	+	none	4
25	Ōgami	-	-	+	none	4
26	Irabu	-	-	+	none	4
27	Hateruma	-	-	+	none	4
28	Hatoma	-	-	+	none	4 (int)
29	Miyara	-	-	+	none	4 (int)
30	Sonai	-	-	+	none	4

*A Data for geographical maps*

	Language	KIN (11)	K~ (12)	who <sup>3</sup> (13)	FRV	PQ vs. Decl
31	Korean	-	-	+	?H+M	3
32	Jeju	-	-	?+	?	3
33	Dagur	+	+	-	none	4
34	Buryat	+	+	-	H+M	4
35	Khamnigan Mongol	+	+	-	?H	4
36	Cyrillic Khalkha	+	+	-	none	4
37	Oirat	+	+	-	H+M	4
38	Shira Yughur	+	+	-	H+M	4
39	Santa	+	+	+	none	4
40	Bonan	+	+	-	?M	4
41	Kangjia	+	-	+	?M+H	4
42	Huzhu Mongghul	+	+	-	none	4
43	Minhe Mangghuer	+	+	-	none	4
44	Standard Mandarin	-	-	-	H	4
45	Urumqi Hui Mandarin	-	-	-	H	4
46	Hezhou	-	-	?	H	4
47	Tangwang	-	-	-	H	4
48	Wutun	-	-	-	none	4
49	Amdo Tibetan, Gonghe	-	-	+	H	4
50	Baima	-	?+	+	?H+M	4
51	Zhongu	-	-	+	none	4
52	Even	-	-	+	?none	4
53	Evenki	-	-	+	none	4
54	Negidal	-	-	+	?	4
55	Kh. Evenki	-	-	+	?none	4
56	Oroqen, Xunke	-	-	+	H	4
57	Solon, Huihe	-	-	+/-	M+H	4
58	Udihe	-	-	+	M+H	4
59	Kilen	-	+	+	M+H	4
60	Nanai	-	+	+	none	4

	Language	KIN (11)	K~ (12)	who <sup>4</sup> (13)	FRV	PQ vs. Decl
61	Ulcha	-	+	+	?	4
62	Uilta	-	+	+	M	4
63	Sibe	-	-	+	H+M	4
64	Salar	+	+	+	H+M	4
65	Kazakh	+	+	+	H+M	4
66	Kyrgyz	+	+	+	H+M	4
67	Tatar (Chinese)	+	+	+	H+M	4
68	Uyghur	+	+	+	H+M	4
69	Tuvan	+	+	-	H+M	4
70	Tofa	+	+	?+	H+M	4
71	Khakas	+	+	+	H+M	4
72	Sarig	+	+	+	H+M	4
73	Altai	+	+	?-	H+M	4
74	Chulym	+	+	+	H+M	4
75	Dolgan	+	+	-	H+M	4
76	Yakut	+	+	+	H+M	4
77	Nganasan	-	+	+	H	3
78	Forest Enets	-	+	+	none	mixed
79	Tundra Nenets	?-	+	-	none	mixed
80	Selkup, Taz	-	+	-	?H+M	4
81	Ket	-	-	-	none	4
82	Kolyma Yukaghir	+	+	+	M	4
83	Tundra Yukaghir	+	+	+	M	4 (int)

<sup>4</sup>The personal interrogative has a different initial consonant from all other interrogatives.



# References

- Aalto, Pentti. 1976. G. J. Ramstedts Onkor-solonisches Wörterverzeichnis I. *Rocznik Orientalistyczny* 38. 31–41.
- Aalto, Pentti. 1979. G. J. Ramstedts Onkor-solonische Sprachmaterialien. *Studia Orientalia* 51(5). 3–13.
- Abdurehim, Esmael. 2014. *The Lopnor dialect of Uyghur. A descriptive analysis* (Publications of the Institute for Asian and African Studies 17). Helsinki: Unigrafia.
- Abondolo, Daniel. 1998. Introduction. In Daniel Abondolo (ed.), *The Uralic languages* (Routledge Language Family Series), 1–42. London: Routledge.
- Acuo Yixiweisa. 2001. Zang han hunheyu “daohua” shulüe. *Yuyan yanjiu* 44(3). 109–126.
- Adams, Douglas Q. 2013. *A dictionary of Tocharian B*. 2nd edn. (Leiden Studies in Indo-European 10). Amsterdam: Rodopi.
- Aikhenvald, Alexandra Y. 2012. The essence of mirativity. *Linguistic Typology* 16. 435–485.
- Aikhenvald, Alexandra Y. & Robert M. W. Dixon. 2001. Introduction. In Alexandra Y. Aikhenvald & Robert M. W. Dixon (eds.), *Areal diffusion and genetic inheritance. Problems in comparative linguistics*, 1–26. Oxford: Oxford University Press.
- Aikio, Ante. 2014. The Uralic-Yukaghir lexical correspondences: Genetic inheritance, language contact or chance resemblance? *Finnisch-Ugrische Forschungen* 62. 7–76.
- Aixinjueluo Yingsheng. 1987a. Manyu yu hanyu de huxiang yingxiang. *Manyu yanjiu* 1. 67–72.
- Aixinjueluo Yingsheng. 1987b. Tantan manyu de jingyu. *Manyu yanjiu* 1. 2–15, 73.
- Åkerman, Vesa. 2012. *Inflection of finite verbs in Mongghul* (SIL Electronic Working Papers 3). Dallas: SIL International.
- Alexander, Richard & Arran Stibbe. 2014. From the analysis of ecological discourse to the ecological analysis of discourse. *Language Sciences* 41. 104–110.
- Allentoft, Morten E., Martin Sikora, Karl-Göran Sjögren, Simon Rasmussen, Morten Rasmussen, Jesper Stenderup, Peter de Barros Damgaard, Hannes Schroeder, Torbjörn Ahlström, Lasse Vinner, Anna-Sapfo Malaspina, Ashot Margaryan, Thomas F. G. Higham, David Chivall, Niels Lynnerup, Lise Harvig, Justyna Baron, Philippe Della Casa, Paweł Dąbrowski, Paul R. Duffy, Alexander V. Ebel, Andrey Epimakhov, Karin Frei, Mirosław Furmanek, Tomasz Gralak, Andrey Gromov, Stanisław Gronkiewicz, Gisela Grupe, Tamás Hajdu, Radosław Jarysz, Valeri Khartanovich, Alexandr Khokhlov, Viktória Kiss, Jan Kolář, Aivar Kriiska, Irena Lasak, Cristina Longhi, George McGlynn, Algimantas Merkevičius, Inga Merkyte, Mait Metspalu, Ruzan Mkrtychyan, Vyacheslav Moiseyev, László Paja, György Pálfi, Dalia Pokutta, Łukasz Pospieszny, T. Douglas Price, Lehti Saag, Mikhail Sablin, Natalia Shishlina,

## References

- Václav Smrčka, Vasiliï I. Soenov, Vajk Szeverényi, Gusztáv Tóth, Synaru V. Trifanova, Liivi Varul, Magdolna Vicze, Levon Yepiskoposyan, Vladislav Zhitenev, Ludovic Orlando, Thomas Sicheritz-Pontén, Søren Brunak, Rasmus Nielsen, Kristian Kristiansen & Eske Willerslev. 2015. Population genomics of Bronze Age Eurasia. *Nature* 522. 167–172.
- Alonso de la Fuente, José Andrés. 2011. Venjukov's [1858] 1862/1868 Nanai materials. *Journal de la Société Finno-Ougrienne* 93. 11–35.
- Alonso de la Fuente, José Andrés. 2017. On internal reconstruction, morphological restructuring and nominal classes in Manchu. *Acta Orientalia Academiae Scientiarum Hung.* 70(3). 347–370.
- Amha, Azeb. 2007. Questioning forms in Zargulla. In Rainer Vogt (ed.), *From beyond the Mediterranean. Akten des 7. Internationalen Semitohamitistenkongresses, Berlin 13. bis 15. September 2004*, vol. 5 (Semitica et Semitohamitica Berolinensia), 197–210. Aachen: Shaker Verlag.
- Amha, Azeb. 2012. Omotic. In Zygmunt Frajzyngier & Erin Shay (eds.), *The Afroasiatic languages* (Cambridge Language Surveys), 423–504. Cambridge: Cambridge University Press.
- An Jun. 1984. Hezheyu gaikuang. *Minzu yuwen* 6. 55–68.
- An Jun. 1986. *Hezheyu jianzhi*. Peking: Minzu chubanshe.
- Anderson, Gregory D. S. 1998. *Xakas* (Languages of the World/Materials 251). Munich: Lincom Europa.
- Anderson, Gregory D. S. 2001. Subject version and object version in Tofa auxiliary constructions. *Turkic Languages* 5. 245–269.
- Anderson, Gregory D. S. 2003. Yeniseic languages from a Siberian areal perspective. *Sprachtypologie und Universalienforschung* 56(1-2). 12–39.
- Anderson, Gregory D. S. 2004. The languages of Central Siberia. Introduction and overview. In Edward Vajda (ed.), *Languages and prehistory of Central Siberia* (Current Issues in Linguistic Theory 262), 1–119. Amsterdam: Benjamins.
- Anderson, Gregory D. S. 2006a. Chukotko-Kamchatkan languages. In Keith Brown (ed.), *Encyclopedia of language & linguistics*, 2nd edn., 415–417. Oxford: Elsevier.
- Anderson, Gregory D. S. 2006b. Ket. In Keith Brown (ed.), *Encyclopedia of language & linguistics*, 2nd edn., 184–186. Oxford: Elsevier.
- Anderson, Gregory D. S. 2006c. Nivkh. In Keith Brown (ed.), *Encyclopedia of language & linguistics*, 2nd edn., 645–647. Oxford: Elsevier.
- Anderson, Gregory D. S. 2006d. Towards a typology of the Siberian linguistic area. In Yaron Matras, April McMahon & Nigel Vincent (eds.), *Linguistic areas. Convergence in historical and typological perspective*, 266–300. Basingstoke: Palgrave MacMillan.
- Anderson, Gregory D. S. 2006e. Yukaghir. In Keith Brown (ed.), *Encyclopedia of language & linguistics*, 2nd edn., 739–741. Oxford: Elsevier.
- Anderson, Gregory D. S. 2010. Perspectives on the global language extinction crisis: The Oklahoma and eastern Siberia language hotspots. *Revue Roumaine de Linguistique* 2. 128–142.

- Anderson, Gregory D. S. 2013. The velar nasal. In Matthew S. Dryer & Martin Haspelmath (eds.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://wals.info/chapter/9>. (Accessed 2015-03-02.)
- Anderson, Gregory D. S. & K. David Harrison. 1999. *Tyvan* (Languages of the World/Materials 257). Munich: Lincom Europa.
- Anderson, Gregory D. S. & K. David Harrison. 2004. Shaman and bear. Siberian prehistory in two Middle Chulym texts. In Edward Vajda (ed.), *Languages and prehistory of Central Siberia* (Current Issues in Linguistic Theory 262), 179–197. Amsterdam: Benjamins.
- Anderson, Gregory D. S. & K. David Harrison. 2006. Ös tili: Towards a comprehensive documentation of Middle and Upper Chulym dialects. *Turkic Languages* 10. 47–71.
- Andvik, Erik E. 2010. *A grammar of Tshangla* (Brill's Tibetan studies library. Languages of the Greater Himalayan region 5(10)). Leiden: Brill.
- Anthony, David W. 2007. *The horse, the wheel and language. How Bronze-Age riders from the Eurasian Steppe shaped the modern world*. Princeton: Princeton University Press.
- Anthony, David W. & Don Ringe. 2015. The Indo-European homeland from linguistic and archaeological perspectives. *Annual Review of Linguistics* 1. 199–219.
- Aoi, Hayato. 2015. Tarama Miyako grammar. In Heinrich Patrick, Shinsho Miyara & Michinori Shimoji (eds.), *Handbook of the Ryukyuan languages. History, structure, and use* (Handbooks of Japanese language and linguistics 11), 405–421. Berlin: de Gruyter Mouton.
- Arakaki, Tomoko. 2003. A grammatical study of Luchuan. Focusing upon the verb forms in Shuri dialect. In Atsuko Izuyama (ed.), *Studies on Luchuan grammar* (Endangered languages of the Pacific Rim A4-024), 163–216. Kyōto: Nakanishi Printing.
- Arakaki, Tomoko. 2010. *Evidentials in the Shuri dialect of Luchuan (Ryukyuan)*. Edinburgh: The University of Edinburgh. (Doctoral dissertation.)
- Arakaki, Tomoko. 2015. The tense-aspect-mood system of the Ryukyuan languages. In Heinrich Patrick, Shinsho Miyara & Michinori Shimoji (eds.), *Handbook of the Ryukyuan languages. History, structure, and use*, 227–252. Berlin: De Gruyter Mouton.
- Araujo, Gabriel Antunes de. 2004. *A grammar of Sabanê. A Nambikwaran language* (LOT 94). Utrecht: LOT.
- Arnheim, Rudolf. 1969. *Visual thinking*. Berkeley: University of California Press.
- Asai, Tōru. 1974. Classification of dialects: Cluster analysis of Ainu dialects. *Bulletin of the Institute for the Study of North Eurasian Cultures* 8. 45–136.
- Aston, W. G. 1904. *A grammar of the Japanese written language*. 3rd edn. London: Luzac & Co.
- Atkine, Victor. 1997. The Evenki language from the Yenisei to Sakhalin. *Senri Ethnological Studies* 44. 109–121.
- Audova, Iris. 1997. On the North-Samoyed interrogative against the background of neighbouring languages. *Fenno-Ugristica* 21. 36–41.
- Austerlitz, Robert. 1956. Gilyak nursery words. *Word* 12(2). 260–279.
- Avrorin, Valentin A. & Boris V. Boldyrev. 1978. *Orochskie teksty i slovar'*. Leningrad: Nauka.

## References

- Avrorin, Valentin A. & Boris V. Boldyrev. 2001. *Grammatika orochskogo jazyka*. Novosibirsk: Izdat.
- Axelsen, Jacob Bock & Susanna Manrubia. 2014. River density and landscape roughness are universal determinants of linguistic diversity. *Proceedings of the Royal Society B: Biological Sciences* 281. 1–8.
- Axelsson, Karin. 2011. A cross-linguistic study of grammatically dependent question tags. Data and theoretical implications. *Studies in Language* 35(4). 793–851.
- Baek, Sangyub. 2016. *Tungusic from the perspective of areal linguistics: Focusing on the Bikin dialect of Udihe*. Sapporo: Hokkaidō University. (Doctoral dissertation.)
- Bai Ping. 2011. Neimonggu eluosiya gaikuang. *Minzu yuwen* 4. 70–80.
- Baitchura, Uzbek. 1979. Intonation and stress in Nanai in connection with those in other Manchu-Tungusic languages according to instrumental-phonetic data. *Zentralasiatische Studien* 13. 281–308.
- Bar, Moshe. 2009. The proactive brain: Memory for predictions. *Philosophical Transactions of the Royal Society B: Biological Sciences* 364. 1235–1243.
- Baranesa, Adrien, Pierre-Yves Oudeyer & Jacqueline Gottlieb. 2015. Eye movements reveal epistemic curiosity in human observers. *Vision Research* 117. 81–90.
- Barsalou, Lawrence W. 2009. Simulation, situated conceptualization, and prediction. *Philosophical Transactions of the Royal Society B: Biological Sciences* 364. 1281–1289.
- Bashir, Elena. 2009. Wakhi. In Gernot Windfuhr (ed.), *The Iranian languages* (Routledge Language Family Series), 825–862. London: Routledge.
- Baskakov, N. S. 1958a. *Altajskij jazyk*. Moscow: Nauka.
- Baskakov, N. S. 1958b. La classification des dialectes de la langue Turque d'Altaï. *Acta Orientalia Academiae Scientiarum Hungaricae* 8(1). 9–15.
- Baskakov, N. S. 1997. Altajskij jazyk. In Viktorija N. Jarceva, Nikolaj Nikolaevič Kazanskij & Andrej Aleksandrovič Kibrik (eds.), *Jazyki mira, vol. 2: Tjurkskiejazyki*, 179–187. Moskva: Nauka.
- Batchelor, John. 1905. *An Ainu-English-Japanese dictionary (including a grammar of the Ainu language)*. 2nd edn. Tōkyō: Methodist Publishing House.
- Baxter, William H. & Laurent Sagart. 2014a. *Old Chinese. A new reconstruction*. Oxford: Oxford University Press.
- Baxter, William H. & Laurent Sagart. 2014b. *Old Chinese reconstructions*. <http://ocbaxtersagart.lsa.it.lsa.umich.edu/>. (Accessed 2016-10-08.)
- Beckwith, Christopher I. 2007. *Koguryo: The language of Japan's continental relatives. An introduction to the historical-comparative study of the Japanese-Koguryoic languages with a preliminary description of archaic northeastern Middle Chinese*. 2nd edn. Leiden: Brill.
- Bellwood, Peter. 2013. Prehistoric migration and the rise of humanity. In Peter Bellwood (ed.), *The global prehistory of human migration*, 1–6. Malden: Wiley Blackwell.
- Bencini, Giulia. 2003. Toward a diachronic typology of yes/no question constructions with particles. *Chicago Linguistic Society* 39. 604–621.
- Bentley, John R. 2008a. *A linguistic history of the forgotten islands*. Folkestone: Global Oriental.

- Bentley, John R. 2008b. The search for the language of Yamatai. *Japanese Language and Literature* 42. 1–43.
- Benzing, Johannes. 1955. *Lamutische Grammatik. Mit Bibliographie, Sprachproben und Glossar* (Akademie der Wissenschaften der Literatur. Veröffentlichungen der orientalischen Kommission 6). Wiesbaden: Harrassowitz.
- Benzing, Johannes. 1956. *Die tungusischen Sprachen: Versuch einer vergleichenden Grammatik* (Akademie der Wissenschaften und der Literatur. Abhandlungen der Geistes- und Sozialwissenschaftlichen Klasse 11). Mainz: Akademie der Wissenschaften und der Literatur.
- Benzing, Johannes. 1985. *Kalmückische Grammatik zum Nachschlagen* (Turcologica 1). Wiesbaden: Harrassowitz.
- Berge, Anna. 2006. Eskimo-Aleut. In Keith Brown (ed.), *Encyclopedia of language & linguistics*, 2nd edn., 219–223. Oxford: Elsevier.
- Berge, Anna. 2010. Origins of linguistic diversity in the Aleutian Islands. *Human Biology* 82(5/6). 557–581.
- Bergen, Benjamin. 2012. *Louder than words. The new science of how the mind makes meaning*. New York: Basic Books.
- Berghäll, Liisa. 2015. *A grammar of Mauwake* (Studies in Diversity Linguistics 4). Berlin: Language Science Press.
- Bergsland, Knut. 1997. *Aleut grammar. Unangam Tunuganaan Achixaasīx̄*. Fairbanks: Alaska Native Language Center.
- Berlyne, Daniel E. 1954. A theory of human curiosity. *British Journal of Psychology* 45(3). 180–191.
- Berlyne, Daniel E. 1960. *Conflict, arousal, and curiosity*. New York: McGraw-Hill Book Company.
- Berlyne, Daniel E. 1978. Curiosity and learning. *Motivation and Emotion* 2(2). 97–175.
- Bhat, D. N. S. 2004. *Pronouns* (Oxford Studies in Typology and Linguistic Theory). Oxford: Oxford University Press.
- Bickel, Balthasar. 2007. Typology in the 21st century: Major current developments. *Linguistic Typology* 11. 239–251.
- Bickel, Balthasar. 2014. Sprachliche Vielfalt im Wechselspiel von Natur und Kultur. In Elvira Glaser, Agnes Kolmer, Martin Meyer & Elisabeth Stark (eds.), *Sprache(n) verstehen. Interdisziplinäre Vortragsreihe der Eidgenössischen Technischen Hochschule Zürich und der Universität Zürich Herbstsemester 2012*, 101–126. Zürich: vdf Hochschulverlag.
- Bickel, Balthasar. 2015. Distributional typology: Statistical inquiries into the dynamics of linguistic diversity. In Bernd Heine & Heiko Narrog (eds.), *The Oxford handbook of linguistic analysis*, 2nd edn., 901–923. Oxford: Oxford University Press.
- Bickel, Balthasar & Johanna Nichols. 2007. Inflectional morphology. In Timothy Shopen (ed.), *Language typology and syntactic description, vol. 3: Grammatical categories and the lexicon*, 2nd edn., 169–240. Cambridge: Cambridge University Press.
- Bickerton, Derek. 2016 [1981]. *Roots of language* (Classics in Linguistics 3). Berlin: Language Science Press.

## References

- Bielmeier, Roland. 1989. Yagnōbī. In Rüdiger Schmitt (ed.), *Compendium linguarum Iranicorum*, 480–488. Wiesbaden: Reichert.
- Bilaniuk, Laada. 2004. A typology of Surzhyk: Mixed Ukrainian-Russian language. *International Journal of Bilingualism* 8(4), 409–425.
- Birjukovich, R. M. 1997. Chulymsko-tjorskij jazyk. In Viktorija N. Jarceva, Nikolaj Nikolaevich Kazanskij & Andrej Aleksandrovich Kibrik (eds.), *Jazyki mira, vol. 2: tjurkskie jazyki*, 491–497. Moskva: Nauka.
- Birtalan, Ágnes. 2003. Oirat. In Juha Janhunen (ed.), *The Mongolic languages* (Routledge Language Family Series), 210–228. London: Routledge.
- Bisang, Walter. 2010. Areal language typology. In Peter Auer & Jürgen E. Schmidt (eds.), *Language and space. An international handbook of linguistic variation* (Handbücher zur Sprach- und Kommunikationswissenschaft 30(1)), 419–440. Berlin: Mouton de Gruyter.
- Bläsing, Uwe. 2003. Kalmuck. In Juha Janhunen (ed.), *The Mongolic languages* (Routledge Language Family Series), 229–247. London: Routledge.
- Blench, Roger & Mark W. Post. 2014. Rethinking Sino-Tibetan phylogeny from the perspective of North East Indian languages. In Nathan W. Hill & Thomas Owen-Smith (eds.), *Trans-Himalayan linguistics. Historical and descriptive linguistics of the Himalayan area* (Trends in Linguistics. Studies and Monographs 266), 71–104. Berlin: De Gruyter Mouton.
- Blust, Robert. 2013. *The Austronesian languages* (Asia-Pacific Linguistics 8). Canberra: Australian National University.
- Bobaljic, Jonathan D. & Susi Wurmbbrand. 2002. Notes on agreement in Itelmen. *Linguistic Discovery* 1(1), 1–27.
- Boeschoten, Hendrik. 1998. Uzbek. In Lars Johanson & Éva Á. Csató (eds.), *The Turkic languages* (Routledge Language Family Series), 357–378. London: Routledge.
- Boeschoten, Hendrik & Marc Vandamme. 1998. Chaghatay. In Lars Johanson & Éva Á. Csató (eds.), *The Turkic languages* (Routledge Language Family Series), 166–178. London: Routledge.
- Bogoras, Waldemar. 1922. Chukchee. In Franz Boas (ed.), *Handbook of American Indian languages, vol. 2*, 631–903. Washington: Government Printing Office.
- Bolinger, Dwight. 1957. *Interrogative structures of American English*. Alabama: University of Alabama Press.
- Bowern, Claire L. 2012. *A grammar of Bardi* (Mouton Grammar Library 57). Berlin: De Gruyter Mouton.
- Bowern, Claire, Patience Epps, Jane Hill & Patrick McConvell. 2016. *Hunter-gatherer language database*. <https://huntergatherer.la.utexas.edu/>. Accessed 2015-07-16.
- Braune, Wilhelm & Frank Heidermanns. 2004. *Gotische Grammatik. Mit Lesestücken und Wörterverzeichnis* (Sammlung kurzer Grammatiken germanischer Dialekte A(1)). Tübingen: May Niemeyer.
- Brosig, Benjamin. 2014. The tense-aspect system of Khorchin Mongolian. In Pirkko Suihkonen & Lindsay J. Whaley (eds.), *On diversity and complexity of languages spo-*

- ken in Europe and North and Central Asia* (Studies in Language Companion Series 164), 3–65. Amsterdam: Benjamins.
- Brown, Lucien. 2011. *Korean honorifics and politeness in second language learning* (Pragmatics & Beyond New Series 206). Amsterdam: Benjamins.
- Brown, Lucien & Jaehoon Yeon. 2015. Varieties of contemporary Korean. In Lucien Brown & Yeon Jaehoon (eds.), *The handbook of Korean linguistics*, 459–476. Malden: Wiley Blackwell.
- Bugaeva, Anna. 2004. *Grammar and folklore texts of the Chitose dialect of Ainu (idiolect of Ito Oda)*. Kyōto: Nakanishi.
- Bugaeva, Anna. 2012. Southern Hokkaido Ainu. In Nicolas Tranter (ed.), *The languages of Japan and Korea* (Routledge Language Family Series), 461–509. London: Routledge.
- Buhe & Liuzhaoxiang. 1982. *Baoanyu jianzhi*. Peking: Minzu chubanshe.
- Bulatova, Nadezhda & Stefano Cotrozzi. 2004. *Ewenkische Erzählungen aus der Insel Sachalin*. Bonn: Verlag für Kultur und Wissenschaft.
- Bulatova, Nadezhda & Lenore Grenoble. 1999. *Evenki* (Languages of the World/Materials 141). Munich: Lincom Europa.
- Cable, Seth. 2007. *The grammar of Q: Q-particles and the nature of wh-fronting, as revealed by the wh-questions of Tlingit*. Cambridge: MIT. (Doctoral dissertation.)
- Campbell, Lyle. 2006. Areal linguistics: A closer scrutiny. In Yaron Matras, April McMahon & Vincent Nigel (eds.), *Linguistic areas. Convergence in historical and typological perspective*, 1–31. Basingstoke: Palgrave MacMillan.
- Carling, Gerd. 2009. *Dictionary and thesaurus of Tocharian A, vol.1: A-ĵ*. Wiesbaden: Harrassowitz.
- Castrén, M. Alexander. 1855. *Wörterverzeichnisse aus den samojedischen Sprachen*. St Petersburg: Buchdruckerei der kaiserlichen Akademie der Wissenschaften.
- Castrén, M. Alexander. 1856. *Grundzüge einer tungusischen Sprachlehre nebst kurzem Wörterverzeichniss*. St Petersburg: Buchdruckerei der kaiserlichen Akademie der Wissenschaften.
- Castrén, M. Alexander. 1857a. *Versuch einer burjätischen Sprachlehre nebst kurzem Wörterverzeichnis*. St Petersburg: Buchdruckerei der kaiserlichen Akademie der Wissenschaften.
- Castrén, M. Alexander. 1857b. *Versuch einer koibalischen und karagassischen Sprachlehre*. St Petersburg: Buchdruckerei der kaiserlichen Akademie der Wissenschaften.
- Castrén, M. Alexander. 1858. *Versuch einer jennissej-ostjakischen und kottischen Sprachlehre*. St Petersburg: Buchdruckerei der kaiserlichen Akademie der Wissenschaften.
- Chae, Heekyung. 2013. Nivufugo higashi saharin hōgen no mukashibanashi tekisuto: Chiharoshutto usagi. *Hoppō jinbun kenkyū* 6. 129–135.
- Chaghanhada. 1991. Mengguyu ke'erqin tuyude yuqici. *Minzu yuwen* 2. 71–73.
- Chaoke D. O. 2006. *Xiandai xiboyu kouyu yanjiu*. Peking: Renmin chubanshe.
- Chaoke D. O. 2007. *Nanmu elunchunyu yanjiu*. Peking: Minzu chubanshe.
- Chaoke D. O. 2009. *Ewenkeyu cankao yufa*. Peking: Zhongguo shehui kexue chubanshe.
- Chaoke D. O. 2014a. *Elunchunyu 366 ju huihua ju*. Peking: Shehui kexue wenxian chubanshe.

## References

- Chaoke D. O. 2014b. *Hezheyu 366 ju huihuaju*. Peking: Shehui kexue wenxian chubanshe.
- Chaoke D. O., Kajia, Aliya & Tami'er. 2014. *Ewenkeyu 366 ju huihuaju*. Peking: Shehui kexue wenxian chubanshe.
- Chaolu Wu. 1994a. Introduction, grammar, and sample sentences for Baoan. *Sino-Platonic Papers* 58. 1–28.
- Chaolu Wu. 1994b. Introduction, grammar, and sample sentences for Dagur. *Sino-Platonic Papers* 56. 1–36.
- Chaolu Wu. 1994c. Introduction, grammar, and sample sentences for Dongxiang. *Sino-Platonic Papers* 55. 1–34.
- Chaolu Wu. 1994d. Introduction, grammar, and sample sentences for Jegün Yogur. *Sino-Platonic Papers* 54. 1–34.
- Chaolu Wu. 1996. *Daur* (Languages of the World/Materials 93). Munich: Lincom Europa.
- Chard, Chester S. 1974. *Northeast Asia in prehistory*. Madison: The University of Wisconsin Press.
- Chen Litong. 2017. Development of Daohua: Social contexts of a Chinese-Tibetan creole. *Asian Highland Perspectives* 44. 122–140.
- Chen Zhaojun, Li Xingzhong, Lü Jinliang, Keith W. Slater, Kevin Stuart, Wang Xianzhen, Wang Yongwei, Wang Zhenlin, Xin Huaizhi, Zhu Meilan, Zhu Shanzhong, Zhu Wenhui & Zhu Yongzhong. 2005. *Folktales of the Minhe Mangghuer* (Languages of the World/Text Library 1). Munich: Lincom Europa.
- Chen Zongzhen. 1982. Xibu yuguyu gaikuang. *Minzu yuwen* 6. 66–78.
- Chen Zongzhen & Yi Liqian. 1986. *Tata'er jianzhi*. Peking: Minzu chubanshe.
- Cheng, Andrew & K. David Harrison. 2014. *Jeju-eo talking dictionary*. [http : / / talkingdictionary.swarthmore.edu/jeju/](http://talkingdictionary.swarthmore.edu/jeju/). (Accessed 2016-06-09.)
- Chien Yuehchen & Sanada Shinji. 2010. Yilan Creole in Taiwan. *Journal of Pidgin and Creole Languages* 25(2). 350–357.
- Chirkova, Katia. 2012. The Qiangic subgroup from an areal perspective: A case study of languages of Muli. *Language and Linguistics* 13(1). 133–170.
- Chisholm, William S. (ed.). 1984. *Interrogativity. A colloquium on the grammar, typology and pragmatics of questions in seven diverse languages, Cleveland, Ohio, October 5th 1981-May 3rd 1982* (Typological Studies in Language 4). Amsterdam: Benjamins.
- Cincius, Vera I. 1949. *Sravnitel'naja fonetika tunguso-man'chzhurskikh jazykov*. Leningrad: Nauka.
- Cincius, Vera I. 1975/77. *Sravnitel'nyj slovar' tunguso-man'chzhurskikh jazykov*. Leningrad: Nauka. 2 vols.
- Cincius, Vera I. 1982. *Negidal'skij jazyk*. Leningrad: Nauka.
- Clark, Larry V. 1998. Chuvash. In Lars Johanson & Éva Á. Csató (eds.), *The Turkic languages* (Routledge Language Family Series), 434–452. London: Routledge.
- Clark, Marybeth. 1985. Asking questions in Hmong and other Southeast Asian languages. *Linguistics of the Tibeto-Burman Area* 8(2). 60–67.
- Comrie, Bernard. 1981. *The languages of the Soviet Union* (Cambridge Language Surveys). Cambridge: Cambridge University Press.

- Comrie, Bernard. 1984. Russian. In Williams S. Chisolm (ed.), *Interrogativity. A colloquium on the grammar, typology and pragmatics of questions in seven diverse languages, Cleveland, Ohio, October 5th 1981-May 3rd 1982* (Typological Studies in Language 4), 7–46. Amsterdam: Benjamins.
- Comrie, Bernard. 2003. Yeniseic linguistics at the crossroads. *Sprachtypologie und Universalienforschung* 56(1/2). 8–11.
- Comrie, Bernard. 2008. The areal typology of Chinese: Between North and Southeast Asia. In Redouane Djamouri, Barbara Meisterernst & Rint Sybesma (eds.), *Chinese Linguistics in Leipzig*, vol. 2 (Chinese Linguistics in Europe), 1–21. Paris: Centre de Recherches Linguistiques sur l'Asie Orientale.
- Comrie, Bernard. 2013. *Is Siberia a linguistic area?* Paper presented at Voices from the Indigenous Siberia, with an emphasis on Yukaghir, 13 December 2013, Amsterdam.
- Couper-Kuhlen, Elizabeth. 2012. Some truths and untruths about final intonation in conversational questions. In Jan P. de Ruiter (ed.), *Questions. Formal, functional and interactional perspectives*, 123–145. Cambridge: Cambridge University Press.
- Creissels, Denis. 2006. Encoding the distinction between location, source and destination. A typological study. In Maya Hickmann & Stéphane Robert (eds.), *Space in languages. Linguistic systems and cognitive categories*, vol. 66 (Typological Studies in Language), 19–28. Amsterdam: Benjamins.
- Croft, William. 2003. *Typology and universals*. 2nd edn. (Cambridge Textbooks in Linguistics). Cambridge: Cambridge University Press.
- Cubberley, Paul. 2002. *Russian*. Cambridge: Cambridge University Press.
- Cysouw, Michael. 2005. *The typology of content interrogatives*. Paper presented at Association for Linguistic Typology, 21–25 July 2005, Padang.
- Cysouw, Michael. 2007. Content interrogatives in Pichis Ashéninka: Corpus study and typological comparison. *International Journal of American Linguistics* 73(2). 133–163.
- Cysouw, Michael & Bernard Comrie. 2013. Some observations on typological features of hunter-gatherer languages. In Balthasar Bickel, Lenore A. Grenoble, David A. Peterson & Alan Timberlake (eds.), *Language typology and historical contingency. In honor of Johanna Nichols* (Typological Studies in Language 104), 383–394. Amsterdam: Benjamins.
- Cysouw, Michael & Olav Hackstein. 2011. *The (in)stability of interrogatives. The case of Indo-European*. Paper presented at 20th International Conference on Historical Linguistics, 25–30 July 2011, Ōsaka.
- Dahl, Östen. 2001. Principles of areal typology. In Martin Haspelmath, Ekkehard König, Wulf Oesterreicher & Wolfgang Raible (eds.), *Language typology and language universals: An international handbook* (Handbücher zur Sprach- und Kommunikationswissenschaft 20(2)), 1456–1470. Berlin: De Gruyter.
- Danielsen, Niels. 1972. *Die Frage. Eine sprachwissenschaftliche Untersuchung* (Det Kongelige Danske Videnskaberne Selskab. Historisk-filosofiske skrifter 7(1)). Copenhagen: Munksgaard.
- Davis, Christopher & Tyler Lau. 2015. Tense, aspect, and mood in Miyara Yaeyaman. In Heinrich Patrick, Shinsho Miyara & Michinori Shimoji (eds.), *Handbook of the*

## References

- Ryukyuan languages. History, structure, and use*, vol. 11 (Handbooks of Japanese Language and Linguistics), 253–297. Berlin: De Gruyter Mouton.
- De Busser, Rik. 2015. The influence of social, cultural, and natural factors on language structure. An overview. In Rik De Busser & Randy J. LaPolla (eds.), *Language structure and environment. Social, cultural, and natural factors*, vol. 6 (Cognitive Linguistic Studies in Cultural Contexts), 1–28. Amsterdam: Benjamins.
- de Reuse, Willem J. 1996. Chukchi, English, and Eskimo: A survey of jargons in the Chukotka Peninsula area. In Ernst Håkon Jahr & Ingvild Broch (eds.), *Language contact in the Arctic. Northern pidgins and contact languages*, vol. 88 (Trends in Linguistics. Studies and Monographs), 47–62. Berlin: Mouton de Gruyter.
- de Reuse, Willem J. 2006. *A practical grammar of the San Carlos Apache language* (Lincom studies in Native American linguistics 51). Munich: Lincom Europa.
- de Roerich, Georges. 1958. *Le parler de l’Amdo*. Rome: Istituto Italiano per il Medio ed Estremo Oriente.
- de Sousa, Hilário. 2015. The far southern Sinitic languages as part of Mainland Southeast Asia. In Nick J. Enfield & Bernard Comrie (eds.), *Languages of Mainland Southeast Asia*, vol. 649 (Pacific Linguistics), 356–439. Berlin: Walter de Gruyter.
- Dediu, Dan. 2011. A Bayesian phylogenetic approach to estimating the stability of linguistic features and the genetic biasing of tone. *Proceedings of the Royal Society B: Biological Sciences* 278. 474–479.
- Dediu, Dan & Stephen C. Levinson. 2013. On the antiquity of language: The reinterpretation of Neandertal linguistic capacities and its consequences. *Frontiers in Psychology* 4. 1–17.
- DeLancey, Scott. 2003. Classical Tibetan. In Graham Thurgood & Randy L. LaPolla (eds.), *The Sino-Tibetan languages* (Routledge Language Family Series), 255–269. London: Routledge.
- Delbrück, B. 1900. *Vergleichende Syntax der indogermanischen Sprachen*. Vol. 3. Strassburg: Tübner.
- Denwood, Philip. 1999. *Tibetan* (London Oriental and African Language Library 3). Amsterdam: Benjamins.
- Derksen, Rick. 2008. *Etymological dictionary of the Slavic inherited lexicon* (Leiden Indo-European Etymological Dictionary Series 4). Leiden: Brill.
- Dewey, John. 1896. The reflex arc concept in psychology. *The Psychological Review* 3(4). 357–370.
- Dewey, John. 1910. *How we think*. Boston: D. C. Heath & Co. Publishers.
- Dexi, Zhu. 1985. Hanyu fangyan de liangzhong fanfu wenju. *Zhongguo yuwen* 184. 10–20.
- Di Cosmo, Nicola. 2006. *The diary of a Manchu soldier in seventeenth-century China. “My service in the army”, by Džengšeo* (Routledge Studies in the Early History of Asia). London: Routledge.
- Dienst, Stefan. 2014. *A grammar of Kulina* (Mouton Grammar Library 66). Berlin: Mouton de Gruyter.

- Diessel, Holger. 2003. The relationship between demonstratives and interrogatives. *Studies in Language* 27(3). 635–655.
- Diessel, Holger. 2006. Demonstratives, joint attention, and the emergence of grammar. *Cognitive Linguistics* 17. 463–489.
- Dik, Simon C. 1997. *The theory of functional grammar*. Vol. 2 (Functional Grammar Series 21). Berlin: Mouton de Gruyter.
- Dillon, James T. 1982. The multidisciplinary study of questioning. *Journal of Educational Psychology* 74(2). 147–165.
- Ding Danqing. 1995. Xinjiang dawo'er jianzhi. *Yuyan yanjiu* 28. 188–195.
- Ding, Picus Sizhi. 2014. *A grammar of Prinmi: Based on the central dialect of northwest Yunnan, China* (Brill's Tibetan Studies Library. Languages of the Greater Himalayan Region 14). Leiden: Brill.
- Dingemans, Mark, Francisco Torreira & Nicholas J. Enfield. 2013. Is "huh?" a universal word? Conversational infrastructure and the convergent evolution of linguistic items. *PlosOne* 8(11). 1–10.
- Dixon, Robert M. W. 2002. *Australian languages. Their nature and development* (Cambridge Language Surveys). Cambridge: Cambridge University Press.
- Dixon, Robert M. W. 2012. *Basic linguistic theory, vol. 3: Further grammatical topics*. Oxford: Oxford University Press.
- Doerfer, Gerhard. 1978a. Classification problems of Tungus. In Gerhard Doerfer & Michael Weiers (eds.), *Tungusica, vol. 1: Beiträge zur Nordasiatischen Kulturgeschichte*, 1–26. Wiesbaden: Harrassowitz.
- Doerfer, Gerhard. 1978b. Urtungusisch \*ö. In Gerhard Doerfer & Michael Weiers (eds.), *Tungusica, vol. 1: Beiträge zur Nordasiatischen Kulturgeschichte*, 66–116. Wiesbaden: Harrassowitz.
- Doerfer, Gerhard. 1985. *Mongolo-Tungusica* (Tungusica 3). Wiesbaden: Harrassowitz.
- Doerfer, Gerhard. 1988. *Grammatik des Chaladsch* (Turcologica 4). Wiesbaden: Otto Harrassowitz.
- Doerfer, Gerhard & Michael Knüppel. 2013. *Armanisches Wörterbuch*. Nordhausen: Verlag Traugott Bautz.
- Donidze, G. I. 1997. Shorskij jazyk. In Viktorija N. Jarceva, Nikolaj Nikolaevich Kazanskij & Andrej Aleksandrovich Kibrik (eds.), *Jazyki mira, vol. 2: Tjurkskie jazyki*, 497–506. Moskva: Nauka.
- Doornenbal, Marius A. 2009. *A grammar of Bantawa. Grammar, paradigm tables, glossary and texts of a Rai language of Eastern Nepal* (LOT 222). Utrecht: LOT.
- Dpal-ldan-bkra-shis, Hu Jun, Hu Ping, Li Dechun, Keith W. Slater, Kevin Stuart, Wang Xianzhen & Zhu Yongzhong. 1996. Language materials of China's Monguor minority: Huzhu Mongghul and Minhe Mangghuer. *Sino-Platonic Papers* 69. i–xi, 1–269.
- Dryer, Matthew S. 2003. Word order in Sino-Tibetan languages from a typological and geographical perspective. In Graham Thurgood & Randy LaPolla (eds.), *Sino-Tibetan languages* (Routledge Language Family Series), 43–55. Richmond: Curzon Press.
- Dryer, Matthew S. 2013a. On the six-way word order typology, again. *Studies in Language* 37(2). 267–301.

## References

- Dryer, Matthew S. 2013b. Order of adjective and noun. In Matthew S. Dryer & Martin Haspelmath (eds.), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://wals.info/chapter/87>. Accessed 2016-07-03.)
- Dryer, Matthew S. 2013c. *Order of adposition and noun phrase*. Matthew S. Dryer & Martin Haspelmath (eds.). Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://wals.info/chapter/85>. (Accessed 2016-07-03.)
- Dryer, Matthew S. 2013d. Order of degree word and adjective. In Matthew S. Dryer & Martin Haspelmath (eds.), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://wals.info/chapter/91>. Accessed 2016-07-03.)
- Dryer, Matthew S. 2013e. Order of demonstrative and noun. In Matthew S. Dryer & Martin Haspelmath (eds.), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://wals.info/chapter/88>. Accessed 2016-07-03.)
- Dryer, Matthew S. 2013f. Order of genitive and noun. In Matthew S. Dryer & Martin Haspelmath (eds.), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://wals.info/chapter/86>. Accessed 2016-07-03.)
- Dryer, Matthew S. 2013g. Order of numeral and noun. In Matthew S. Dryer & Martin Haspelmath (eds.), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://wals.info/chapter/89>. Accessed 2016-07-03.)
- Dryer, Matthew S. 2013h. Order of relative clause and noun. In Matthew S. Dryer & Martin Haspelmath (eds.), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://wals.info/chapter/90>. Accessed 2016-07-03.)
- Dryer, Matthew S. 2013i. Order of subject, object and verb. In Matthew S. Dryer & Martin Haspelmath (eds.), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://wals.info/chapter/81>. Accessed 2016-07-03.)
- Dryer, Matthew S. 2013j. Polar questions. In Matthew S. Dryer & Martin Haspelmath (eds.), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://wals.info/chapter/116>. (Accessed 2016-01-23.)
- Dryer, Matthew S. 2013k. Position of interrogative phrases in content questions. In Matthew S. Dryer & Martin Haspelmath (eds.), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://wals.info/chapter/93>. Accessed 2016-07-03.)
- Dryer, Matthew S. 2013l. Position of polar question particles. In Matthew S. Dryer & Martin Haspelmath (eds.), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://wals.info/chapter/92>. (Accessed 2016-09-21.)

- Dryer, Matthew S. 2013m. Prefixing vs. suffixing in inflectional morphology. In Matthew S. Dryer & Martin Haspelmath (eds.), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://wals.info/chapter/26>. Accessed 2016-09-16.)
- Dryer, Matthew S. & Martin Haspelmath (eds.). 2013. *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://wals.info/>. (Accessed 2015-12-09.)
- Duggan, Ana T., Mark Whitten, Victor Wiebe, Michael H. Crawford, Anne Butthof, Victor A. Spitsyn, Sergey Makarov, Innokentiy Novgorodov, Vladimir Osakovsky & Brigitte Pakendorf. 2013. Investigating the prehistory of Tungusic peoples of Siberia and the Amur-Ussuri region with complete mtDNA genome sequences and Y-chromosomal markers. *Plos One* 8(12). 1–19.
- Duncker, Karl & Isadore Krechevsky. 1939. On solution-achievement. *Psychological Review* 46(2). 176–185.
- Dunkel, George E. 2014. *Lexikon der indogermanischen Partikeln und Pronominalstämme* (Indogermanische Bibliothek 2: Reihe: Wörterbücher). Heidelberg: Winter.
- Dunn, Michael. 1999. *A grammar of Chukchi*. Canberra: The Australian National University. (Doctoral dissertation).
- Dunn, Michael. 2000. Chukchi women's language: A historical-comparative perspective. *Anthropological Linguistics* 42(3). 305–328.
- Dwyer, Arienne M. 1995. From the northwest China sprachbund: Xúnhuà Chinese dialect data. *Yuen Ren Society Treasury of Chinese Dialect Data* 1. 143–182.
- Ebata, Fuyuki. 2011. Sakha (Yakut). In Yasuhiro Yamakoshi (ed.), *Grammatical sketches from the field*, 179–212. Tōkyō: Research Institute for Languages, Cultures of Asia & Africa.
- Eberhard, Wolfram. 1965. Chinese regional stereotypes. *Asian Survey* 5(12). 596–608.
- Ebihara, Shiho. 2009. Text from Amdo Tibetan. “Little frog as god's son” (Zarək ʎasi). *Asian and African Languages and Linguistics* 4. 149–168.
- Ebihara, Shiho. 2011. Amdo Tibetan. In Yasuhiro Yamakoshi (ed.), *Grammatical sketches from the field*, 41–78. Tōkyō: Research Institute for Languages, Cultures of Asia & Africa.
- Ebihara, Shiho. 2013. Preliminary field report on dPa'ris dialect of Amdo Tibetan. *Journal of Research Institute* 49. 149–161.
- Ebina, Daisuke. 2011. Cusco Quechua. In Yasuhiro Yamakoshi (ed.), *Grammatical sketches from the field*, 1–39. Tōkyō: Research Institute for Languages, Cultures of Asia & Africa.
- Emmerick, Ronald E. 1985. *The Tumshuqese Karmavācanā text* (Akademie der Wissenschaften und der Literatur. Abhandlungen der Geistes- und Sozialwissenschaftlichen Klasse 2). Mainz: Akademie der Wissenschaften und der Literatur.
- Emmerick, Ronald E. 2009. Khotanese and Tumshuqese. In Gernot Windfuhr (ed.), *The Iranian languages* (Routledge Language Family Series), 377–415. London: Routledge.
- Enfield, Nicholas J. 2010. Questions and responses in Lao. *Journal of Pragmatics* 42. 2649–2665.

## References

- Enfield, Nicholas J. 2014. *Natural causes of language. Frames, biases, and cultural transmission* (Conceptual Foundations of Language Science 1). Berlin: Language Science Press.
- Enfield, Nicholas J. & Bernard Comrie. 2015. Mainland Southeast Asian languages. In Nicholas J. Enfield & Bernard Comrie (eds.), *Languages of Mainland Southeast Asia* (Pacific Linguistics 649), 1–27. Berlin: Walter de Gruyter.
- Enfield, Nicholas J., Tanya Stivers & Stephen C. Levinson. 2010. Question–response sequences in conversation across ten languages: An introduction. *Journal of Pragmatics* 42. 2615–2619.
- Enhebatu, Merden. 1995. *Manyu kouyu yanjiu*. Hohhot: Neimonggu daxue chubanshe.
- Enrico, John. 2004. Toward Proto-Na-Dene. *Anthropological Linguistics* 46(3). 229–302.
- Epps, Patience. 2008. *A grammar of Hup* (Mouton Grammar Library 43). Berlin: Mouton de Gruyter.
- Epps, Patience, John Huehnergard & Na’ama Pat-El. 2013. Introduction: Contact among genetically related languages. *Journal of Language Contact* 6. 209–219.
- Erdal, Marcel. 1998. Old Turkic. In Lars Johanson & Éva Á. Csató (eds.), *The Turkic languages* (Routledge Language Family Series), 138–157. London: Routledge.
- Erdal, Marcel. 2004. *A grammar of Old Turkic* (Handbuch der Orientalistik 8(3)). Leiden: Brill.
- Erdal, Marcel, Irina Nevskaya, Hans Nugteren & Monika Rind-Pawłowski (eds.). 2013. *Handbuch des Tschalkantürkischen. Teil 1: Texte und Glossar* (Turcologica 97). Wiesbaden: Harrassowitz.
- Evans, Nicholas & Stephen C. Levinson. 2009. The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences* 32. 429–492.
- Evans, Vyvyan. 2012. Cognitive linguistics. *WIREs Cogn Sci* 3. 129–141.
- Everett, Caleb, Damián E. Blasi & Seán G. Roberts. 2015. Climate, vocal folds, and tonal languages: Connecting the physiological and geographic dots. *Proceedings of the National Academy of Sciences* 112(5). 1322–1327.
- Everett, Caleb, Damián E. Blasi & Seán G. Roberts. 2016. Language evolution and climate: The case of desiccation and tone. *Journal of Language Evolution* 1(1). 33–46.
- Everett, D. L. & B. Kern. 2007. *Wari’: The Pacaas Novos language of western Brazil* (Descriptive Grammars). London: Routledge.
- Faehndrich, Burgel R. M. 2007. *Sketch grammar of the Karlong variety of Mongghul, and dialectal survey of Mongghul*. Honolulu: University of Hawai’i. (Doctoral dissertation.)
- Field, Kenneth L. 1997. *A grammatical overview of Santa Mongolian*. Oakland: University of California. (Doctoral dissertation.)
- Flegontov, Pavel, Piya Changmai, Anastassiya Zidkova, Maria D. Logacheva, N. Ezgi Altınışık, Olga Flegontova, Mikhail S. Gelfand, Evgeny S. Gerasimov, Ekaterina E. Khrameeva, Olga P. Konovalova, Tatiana Neretina, Yuri V. Nikolsky, George Starostin, Vita V. Stepanova, Igor V. Travinsky, Martin Tříška, Petr Tříška & Tatiana V. Tatarinova. 2016. Genomic study of the Ket: A Paleo-Eskimo-related ethnic group with significant Ancient North Eurasian ancestry. *Scientific Reports* 6. 1–12.

- Foley, William A. 2000. The languages of New Guinea. *Annual Review of Anthropology* 29. 357–404.
- Forsyth, James. 1992. *A history of the peoples of Siberia. Russia's North Asian colony 1581–1990*. Cambridge: Cambridge University Press.
- Fortescue, Michael. 1998. *Language relations across Bering Strait. Reappraising the archaeological and linguistic evidence* (Open Linguistics). London: Cassell.
- Fortescue, Michael. 2003. Diachronic typology and the genealogical unity of Chukotko-Kamchatkan. *Linguistic Typology* 7. 51–88.
- Fortescue, Michael. 2004. How far west into Asia have Eskimo languages been spoken, and which ones? *Études/Inuit/Studies* 28(2). 159–183.
- Fortescue, Michael. 2005. *Comparative Chukotko-Kamchatkan dictionary* (Trends in Linguistics. Documentation 23). Berlin: Mouton de Gruyter.
- Fortescue, Michael. 2011. The relationship of Nivkh to Chukotko-Kamchatkan revisited. *Lingua* 121. 1359–1376.
- Fortescue, Michael. 2013. North America: Eskimo-Aleut linguistic history. In Peter Bellwood (ed.), *The global prehistory of human migration*, 340–345. Malden: Wiley Blackwell.
- Fortescue, Michael. 2016. *Comparative Nivkh dictionary* (Languages of the World/Dictionaries 62). Munich: Lincom Europa.
- Fortescue, Michael, Steven Jacobson & Lawrence Kaplan. 2010. *Comparative Eskimo dictionary: With Aleut cognates*. 2nd edn. Fairbanks: Alaska Native Language Center.
- Fortson, Benjamin W. 2006. Proto-Indo-European syntax. In Keith Brown (ed.), *Encyclopedia of language & linguistics*, 2nd edn., 228–235. Oxford: Elsevier.
- Fortson, Benjamin W. 2010. *Indo-European language and culture. An introduction* (Blackwell Textbooks in Linguistics 30). Malden: Whaley-Blackwell.
- Fountain, Amy. 2008. Introduction to Navajo language studies. *Coyote Papers* 16. 9–34.
- Frellesvig, Bjarke & John Whitman. 2004. The vowels of Proto-Japanese. *Japanese Language and Literature* 38(2). 281–299.
- Fried, Robert Wayne. 2010. *A grammar of Bao'an Tu, a Mongolic language of northwest China*. Buffalo: University of Buffalo. (Doctoral dissertation.)
- Funk, Dimitrij A. 2000. Materialien zur gegenwärtigen Situation der Uilta-Sprache. *Zentralasiatische Studien* 30. 143–152.
- Gallese, Vittorio. 2009. Mirror neurons, embodied simulation, and the neural basis of social identification. *Psychoanalytic Dialogues* 19(5). 519–536.
- Gao Erqiang. 1985. *Tajikeyu jianzhi*. Peking: Minzu chubanshe.
- Gáspár, Csaba. 2006. *Darkhat* (Languages of the World/Materials 419). Munich: Lincom Europa.
- Geiger, Wilhelm. 1901. Die Pāmir-Dialekte, Anhang: Über das Yaghñōbī. In Wilhelm Geiger & Ernst Kuhn (eds.), *Grundriss der iranischen Philologie*, vol. 1(2), 288–344. Strassburg: Trübner.
- Geng Shimin. 2011. Gudai hetian-saiyu gaikuang. *Yuyan yu fanyi* 1. 24–28. , 71.
- Geng Shimin. 2012. Gudai yanqiyu (jiazhong tuhuoluoyu) gaiyao. *Yuyan yu fanyi* 2. 31–39.

## References

- Geng Shimin & Li Zengxiang. 1985. *Hasakeyu jianzhi*. Peking: Minzu chubanshe.
- Georg, Stefan. 2003a. Mongghul. In Juha Janhunen (ed.), *The Mongolic languages* (Routledge Language Family Series), 286–306. London: Routledge.
- Georg, Stefan. 2003b. Ordos. In Juha Janhunen (ed.), *The Mongolic languages* (Routledge Language Family Series), 193–209. London: Routledge.
- Georg, Stefan. 2004. Unreclassifying Tungusic. In Carsten Naehrer (ed.), *Proceedings of the first international conference on Manchu-Tungus studies (Bonn, August 28 - September 1, 2000)* (Tunguso Sibirica 9), 45–57. Wiesbaden: Harrassowitz. vol. 2.
- Georg, Stefan. 2007. *A descriptive Grammar of Ket (Yenisei-Ostyak)*, vol. 1: Introduction, phonology, morphology (Languages of Asia 1). Folkestone: Global Oriental.
- Georg, Stefan. 2008. Yeniseic languages and the Siberian linguistic area. *Studies in Slavic and General Linguistics* 33. 151–168.
- Georg, Stefan & Alexander P. Volodin. 1999. *Die itelmenische Sprache. Grammatik und Texte* (Tunguso Sibirica 5). Wiesbaden: Harrassowitz.
- Gibson, Eleanor J. 1988. Exploratory behavior in the development of perceiving, acting, and the acquiring of knowledge. *Annual Review of Psychology* 39. 1–42.
- Gibson, James J. 1960. The concept of stimulus in Psychology. *The American Psychologist* 15. 694–703.
- Gibson, James J. 1979. *The ecological approach to visual perception*. New York: Psychology Press.
- Girfanova, Albina H. 2002. *Udeghe* (Languages of the World/Materials 255). Munich: Lincom Europa.
- Glenberg, Arthur M. 1997. What memory is for. *Behavioral and Brain Sciences* 20. 1–55.
- Göksel, Aslı & Celia Kerslake. 2005. *Turkish. A comprehensive grammar* (Comprehensive Grammars). London: Routledge.
- Golden, Peter B. 1998. The Turkic peoples: A historical sketch. In Lars Johanson & Éva Á. Csató (eds.), *The Turkic languages* (Routledge Language Family Series), 16–29. London: Routledge.
- Golovko, Evgenij V. 2003. A case of nongenetic development in the Arctic area: The contribution of Aleut and Russian to the formation of Copper Island Aleut. In Ernst Häkon Jahr & Ingvild Broch (eds.), *Language contact in the Arctic. Northern pidgins and contact languages* (Trends in Linguistics. Studies and Monographs), 63–77. Berlin: Mouton de Gruyter.
- Golovko, Evgenij V. & Nikolai Vakhtin. 1990. Aleut in contact: The CIA enigma. *Acta Linguistica Hafniensia* 72. 97–125.
- Gong Hwang-Cherng. 2003. Tangut. In Graham Thurgood & Randy L. LaPolla (eds.), *The Sino-Tibetan languages* (Routledge Language Family Series), 602–620. London: Routledge.
- Gorelova, Liliya M. 2002. *Manchu grammar* (Handbuch der Orientalistik 8(7)). Leiden: Brill.
- Graczyk, Randolph. 2007. *A grammar of Crow*. Lincoln: University of Nebraska Press.

- Graesser, Arthur C. 1985. An introduction to the study of questioning. In Arthur C. Graesser & John B. Black (eds.), *The psychology of questions*, 1–14. Hillsdale: Lawrence Erlbaum.
- Greenberg, Joseph H. 2000. *Indo-European and its closest relatives. The Eurasiatic language family, vol 1: Grammar*. Stanford: Stanford University Press.
- Grube, Wilhelm. 1892. *Giljakisches Wörterverzeichnis nach den Originalaufzeichnungen von L. V. Schrenck und P. V. Glehn* (Reisen und Forschungen im Amur-Lande 3(2.1)). St Petersburg: Kaiserliche Akademie der Wissenschaften.
- Grube, Wilhelm. 1900. *Goldisch-deutsches Wörterverzeichnis mit vergleichender Berücksichtigung der übrigen tungusischen Dialekte* (Reisen und Forschungen im Amur-Lande 3(2.s)). Sankt Petersburg: Bei den Commissionaren der Kaiserlichen Akademie der Wissenschaften.
- Gruzdeva, Ekaterina. 1998. *Nivkh* (Languages of the World/Materials 111). Munich: Lincom Europa.
- Gruzdeva, Ekaterina. 2008. How to orient oneself on Sakhalin: A guide to Nivkh locational terms. *Studies in Slavic and General Linguistics* 33. 169–188.
- Gusev, Valentin. 2015a. Negation in Nganasan. In Matti Miestamo, Anne Tamm & Beáta Wagner-Nagy (eds.), *Negation in Uralic languages* (Typological Studies in Language 108), 102–132. Amsterdam: Benjamins.
- Gusev, Valentin. 2015b. Some parallels in grammar between Nivkh and Tungusic languages. *Journal of the Center for Northern Humanities* 8. 63–75.
- Haak, Wolfgang, Iosif Lazaridis, Nick Patterson, Nadin Rohland, Swapan Mallick, Bastien Llamas, Guido Brandt, Susanne Nordenfelt, Eadaoin Harney, Kristin Stewardson, Fu Qiaomei, Alissa Mittnik, Eszter Bánffy, Christos Economou, Michael Francken, Susanne Friederich, Rafael Garrido Pena, Fredrik Hallgren, Valery Khartanovich, Aleksandr Khokhlov, Michael Kunst, Pavel Kuznetsov, Harald Meller, Oleg Mochalov, Vyacheslav Moiseyev, Nicole Nicklisch, Sandra L. Pichler, Roberto Risch, Manuel A. Rojo Guerra, Christina Roth, Anna Szécsényi-Nagy, Joachim Wahl, Matthias Meyer, Johannes Krause, Dorcas Brown, David W. Anthony, Alan Cooper, Kurt Werner Alt & David Reich. 2015. Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature* 522. 207–211.
- Hackstein, Olav. 2004. From discourse to syntax: The case of compound interrogatives in Indo-European and beyond. In Karlene Jones-Bley, Martin E. Huld, Angela della Volpe & Miriam Robbins Dexter (eds.), *Proceedings of the fifteenth annual UCLA Indo-European conference*, 257–298. Washington: Institute for the Study of Man.
- Hackstein, Olav. 2013. Polar questions and non-headed conditionals in cross-linguistic and historical perspective. In Benjamin Slade & Chen Shu-Fen (eds.), *Grammatica et verba. Glamor and verve. Studies in South Asian, historical, and Indo-European linguistics in Honor of Hans Henrich Hock on the occasion of his seventy-fifth birthday*, 99–116. Ann Arbor: Beech Stave Press.
- Hagège, Claude. 2008. Towards a typology of interrogative verbs. *Linguistic Typology* 12. 1–44.

## References

- Hahn, Reinhard F. 1991. An annotated sample of Ili Turki. *Acta Orientalia Academiae Scientiarum Hungaricae* 45(1). 31–53.
- Hahn, Reinhard F. 1998. Yellow Uyghur and Salar. In Lars Johanson & Éva Á. Csató (eds.), *The Turkic languages* (Routledge Language Family Series), 397–402. London: Routledge.
- Hai Feng. 2002. Dongganyu gaikuang. *Minzu yuwen* 1. 70–81.
- Hajdú, Péter. 1988. Die samojedischen Sprachen. In Denis Sinor (ed.), *The Uralic languages* (Handbuch der Orientalistik 8(1)), 3–40. Leiden: Brill.
- Häkkinen, Jaakko. 2012. Early contacts between Uralic and Yukaghir. *Mémoires de la Société Finno-Ougrienne* 264. 91–101.
- Haller, Felix. 2004. *Dialekt und Erzählungen von Themchen: Sprachwissenschaftliche Beschreibung eines Nomadendialektes aus Nord-Amdo* (Beiträge zur tibetischen Erzählforschung 14). Bonn: VGH-Wissenschaftsverlag.
- Hammarström, Harald, Robert Forkel, Martin Haspelmath & Sebastian Bank. 2016. *Glottolog 2.7*. Jena. <http://glottolog.org>. (Accessed 2016-04-06.)
- Han Youfeng & Meng Shuxian. 1993. *Elunchun hanyu duizhao duben*. Peking: Zhongyong minzuyuan chubanshe.
- Harrison, K. David. 2005. A Tuvan hero tale, with commentary, morphemic analysis, and translation. *Journal of the American Oriental Society* 125(1). 1–30.
- Harrison, K. David & Gregory D. S. Anderson. 2003. Middle Chulym: Theoretical aspects, recent fieldwork and current state. *Turkic languages* 7. 245–256.
- Hasegawa, Yoko. 2015. *Japanese. A linguistic introduction*. Cambridge: Cambridge University Press.
- Hashimoto, Montaro J. 1986. The Altaicization of northern Chinese. In John McCoy & Timothy Light (eds.), *Contributions to Sino-Tibetan studies* (Cornell Linguistic Contributions 5), 76–98. Leiden: Brill.
- Hasibate'er. 2016. *Aoluguya fangyan yanjiu*. Peking: Minzu chubanshe.
- Haspelmath, Martin. 2007. Coordination. In Timothy Shopen (ed.), *Language typology and syntactic description, vol. 3: Grammatical categories and the lexicon*, 2nd edn., 1–51. Cambridge: Cambridge University Press.
- Haspelmath, Martin & the APiCS Consortium. 2013a. Polar questions. In Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds.), *Atlas of pidgin and creole language structures online*, 1–51. Leipzig: MPI for Evolutionary Anthropology. <http://apics-online.info/parameters/103%5C#2/30.3/10.0>. (Accessed 2017-07-06.)
- Haspelmath, Martin & the APiCS Consortium. 2013b. Position of interrogative phrases in content questions. In Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds.), *Atlas of pidgin and creole language structures online*, 1–51. Leipzig: MPI for Evolutionary Anthropology. <http://apics-online.info/parameters/12%5C#2/30.3/10.0>. (Accessed 2017-07-06.)
- Hauer, Erich. 2007. *Handwörterbuch der Mandschusprache*. 2nd edn. Wiesbaden: Harrassowitz.

- Hayashi, Makoto. 2010a. An overview of the question–response system in Japanese. *Journal of Pragmatics* 42. 2685–2702.
- Hayashi, Yuka. 2010b. Ikema (Miyako Ryukyuan). In Thomas Pellard & Michinori Shimoji (eds.), *An introduction to Ryukyuan languages*, 167–188. Tōkyō: Research Institute for Languages, Cultures of Asia & Africa.
- Heggarty, Paul. 2013. Europe and western Asia. Indo-European linguistic history. In Peter Bellwood (ed.), *The global prehistory of human migration*, 157–167. Chichester: Wiley-Blackwell.
- Heggarty, Paul & Colin Renfrew. 2014a. East Asia: Languages. In Colin Renfrew & Paul Bahn (eds.), *The Cambridge world prehistory*, 870–896. Cambridge: Cambridge University Press. 3 vols.
- Heggarty, Paul & Colin Renfrew. 2014b. Introduction: Languages. In Colin Renfrew & Paul Bahn (eds.), *The Cambridge world prehistory*, 19–44. Cambridge: Cambridge University Press. 3 vols.
- Heine, Bernd, Ulrike Claudi & Friederike Hünemeyer. 1991. *Grammaticalization. A conceptual framework*. Chicago: University of Chicago Press.
- Heine, Bernd & Tania Kuteva. 2002. *World lexicon of grammaticalization*. Cambridge: Cambridge University Press.
- Helinski, Eugen. 1997. *Die Matorische Sprache. Wörterverzeichnis, Grundzüge der Grammatik, Sprachgeschichte* (Studia Uralo-Altaica 41). Szeged: JATE Finnugor Tanszék.
- Helinski, Eugen. 1998. Nganasan. In Daniel Abandolo (ed.), *The Uralic languages* (Routledge Language Family Series), 480–515. London: Routledge.
- Hellenthal, Anne-Christie. 2010. *A grammar of Sheko* (LOT 258). Utrecht: LOT.
- Hengeveld, Kees, Maria Luiza Braga, Elisiene de Melo Barbosa, Jaqueline Silveira Coriolano, Juliana Jezuino da Costa, Mariana de Souza Martins, Diego Leite de Oliveira, Vinicius Maciel de Oliveira, Luana Gomes Pereira, Liliane Santana, Cassiano Luiz do Carmo Santos & Viviane dos Ramos Soares. 2012. Semantic categories in the indigenous languages of Brazil. *Functions of Language* 19(1). 38–57.
- Henrich, Joseph, Steven J. Heine & Ara Norenzayan. 2010. The weirdest people in the world? *Behavioral and Brain Sciences* 33. 61–83.
- Hinds, John. 1984. Japanese. In Williams S. Chisolm (ed.), *Interrogativity. A colloquium on the grammar, typology and pragmatics of questions in seven diverse languages*, Cleveland, Ohio, October 5th 1981–May 3rd 1982 (Typological Studies in Language 4), 145–188. Amsterdam: Benjamins.
- Hirofumi, Matsumura & Marc Oxenham. 2013. Eastern Asia: Japan: Human biology. In Peter Bellwood (ed.), *The global prehistory of human migration* (Routledge Language Family Series), 217–223. Malden: Wiley Blackwell.
- Hodges, Bert H. 2009. Ecological pragmatics. Values, dialogical arrays, complexity, and caring. *Pragmatics & Cognition* 17(3). 628–652.
- Hoeks, John C. J., Laurie A. Stowe, Petra Hendriks & Harm Brouwer. 2013. Questions left unanswered: How the brain responds to missing information. *PlosONE* 8(10). 1–9.

## References

- Hölzl, Andreas. 2012. Negation und die Stellung des Mandschurischen innerhalb der tungusischen Sprachen. In Daniel Leese (ed.), *Hou Han(s') Shu: Festschrift zum 50. Geburtstag von Hans van Ess*, 35–67. Munich: Institut für Sinologie.
- Hölzl, Andreas. 2014a. Review of Zikmundová, Veronika. 2013. *Spoken Sibe. Morphology of the inflected parts of speech*. Karolinum Press: Prague. 231 pp. ISBN 978-80-246-2103-6. *Studies in Language* 38. 211–219.
- Hölzl, Andreas. 2014b. “What’s your name?” in *Tungusic*. Paper presented at The 8th International Conference on Construction Grammar, 03–06 September 2014, Osnabrück.
- Hölzl, Andreas. 2015a. A typology of negation in Tungusic. *Studies in Language* 39(1). 117–157.
- Hölzl, Andreas. 2015b. Ecological metaphors. In Andreas Hölzl, Matthias Klumm, Mara Matičević, Thomas Scharinger, Johannes Ungelenk & Nora Zapf (eds.), *Politik der Metapher*, 173–191. Würzburg: Königshausen & Neumann.
- Hölzl, Andreas. 2015c. *Hierarchy or conceptual space: Interogative words in Tungusic*. Paper presented at Polish Cognitive Linguistics Association, 24–26 September 2015, Lublin.
- Hölzl, Andreas. 2015d. *The hunt for polar questions*. Paper presented at The 13th International Cognitive Linguistics Conference, 20–25 July 2015, Newcastle.
- Hölzl, Andreas. 2015e. *Towards a conceptual space of questions*. Paper presented at the 48th Annual Meeting of the Societas Linguistica Europaea, 02–05 September 2015, Leiden.
- Hölzl, Andreas. 2016a. *A typology of question marking in Northeast Asia*. Paper presented at the 6th UK Cognitive Linguistics Conference, 2016.07.19–22, Bangor.
- Hölzl, Andreas. 2016b. Towards a new typology of questions. *Bavarian Working Papers in Linguistics* 5. 17–28.
- Hölzl, Andreas. 2017a. *Kilen. Synchronic and diachronic profile of a mixed language*. Paper presented at the 24. LIPP-Symposium, 21–23 June 2017, Munich.
- Hölzl, Andreas. 2017b. New evidence on Para-Mongolic numerals. *Journal de la Société Finno-Ougrienne* 96. 21–37.
- Hölzl, Andreas. 2017c. *The simplification and complexification of interrogative systems: The case of Tungusic*. Paper presented at Language change for the worse, 25–28 May 2017, Munich.
- Hölzl, Andreas. 2018a. 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). Amsterdam: Benjamins.
- Hölzl, Andreas. 2018b. Udi, Udihe, and the language(s) of the Kyakala. *International Journal of Diachronic Linguistics and Linguistic Reconstruction* 15. 111–146.
- Hong Shi, Qi Xuebin, Hua Zhong, Peng Yi, Zhang Xiaoming, Ma Runlin Z. & Su Bing. 2013. Genetic evidence of an East Asian origin and Paleolithic northward migration of Y-chromosome Haplogroup N. *PlosOne* 8(6). 1–9.
- Hoymann, Gertie. 2010. Questions and responses in †Akhoe Haillom. *Journal of Pragmatics* 42. 2726–2740.
- Hu Zengyi. 2001. *Elunchunyu yanjiu*. Peking: Renmin chubanshe.

- Hu Zengyi & Chaoke D. O. 1986. *Ewenkeyu jianzhi*. Peking: Minzu chubanshe.
- Hu Zhenhua. 1979. Ke'erkezizu yuyan gaikuang. *Xinjiang daxue xuebao* 4. 74–91.
- Hu Zhenhua. 1986. *Ke'erkezizu jianzhi*. Peking: Minzu chubanshe.
- Hu Zhenhua. 1996. Youguan heilongjiangsheng ke'erkezizu de bufen yuyan cailiao. *Minzuyuwen* 5. 70–80.
- Hu Zhenhua & Guy Imart. 1987. Fu-yü Girgīs: A tentative description of the easternmost Turkic language. *Papers on Inner Asia* 8. 1–61.
- Huang, Lillian M. 1996. Interrogative constructions in Mayrinax Atayal. *Bulletin of National Taiwan Normal University* 41. 263–296.
- Huang, Lillian M., Elisabeth Zeitoun, Yeh Marie M., Chang Anna H. & Wu Joy J. 1999. Interrogative constructions in some Formosan languages. In Yin Yuen-mei, Yang I-li & Chan Hui-chen (eds.), *Chinese languages and linguistics, vol. 5: Interactions in Language*, 639–680. Taipei: Institute of Linguistics.
- Huddleston, Rodney. 1994. The contrast between interrogatives and questions. *Journal of Linguistics* 30(2). 411–439.
- Hugjiltu, Wu. 2003. Bonan. In Juha Janhunen (ed.), *The Mongolic languages* (Routledge Language Family Series), 325–345. London: Routledge.
- Hyman, Larry M. & William R. Leben. 2000. Suprasegmental processes. In Geert Booij, Christian Lehmann & Joachim Mugdan (eds.), *Morphology: An international handbook on inflection and word-formation* (Handbücher zur Sprach- und Kommunikationswissenschaft 17(1)), 587–594. Berlin: De Gruyter Mouton.
- Idiatov, Dmitry. 2007. *A typology of non-selective interrogative pronominals*. Antwerp: Universiteit Antwerpen. (Doctoral dissertation.)
- Idiatov, Dmitry & Johan van der Auwera. 2004. On interrogative pro-verbs. In Ileana Comorovski & Manfred Krifka (eds.), *Proceedings of the workshop on the syntax, semantics and pragmatics of questions, ESSLLI 16, August 9–13, 2004, Nancy, France*, 17–23. Nancy: The European Association for Logic, Language & Information.
- Iggesen, Oliver A. 2013. Number of cases. In Matthew S. Dryer & Martin Haspelmath (eds.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://wals.info/chapter/49>. (Accessed 2016-07-04.)
- Ikegami, Jirō. 1958. Orok texts. *Memoirs of the Research Department of the Toyo Bunko* 17. 85–95.
- Ikegami, Jirō. 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.
- Ikegami, Jirō. 1985. The category of person in Tungus: Its representation in the indicative forms of verbs. *Gengō Kenkyū* 88. 86–96.
- Ikegami, Jirō. 1997. *Uirutago jiten*. Sapporo: Hokkaidō University Press.
- Ikegami, Jirō. 1999. *Manshūgo kenkyū*. Tōkyō: Kyūko Shoin.
- Ikegami, Jirō. 2002. *Uiruta kōtō bungei genbunshū*. 2nd edn. (Publications on Tungus languages and cultures 16). Kyōto: Nakanishi.

## References

- Iksop, Lee & Robert S. Ramsey. 2000. *The Korean language* (SUNY series in Korean studies). Albany: State University of New York Press.
- Ivanovskij, A. O. 1982 [1894]. *Mandjurica I. Specimens of the Solon and Dagur languages* (Debter, Deb-Ther, Debtelin: Materials for Central Asiatic and Altaic Studies 2). Budapest: Akadémiai Kaidó.
- Izuyama, Atsuko. 2003. The grammar of Ishigaki Miyara dialect in Luchuan. In Atsuko Izuyama (ed.), *Studies on Luchuan grammar* (Endangered languages of the Pacific Rim), 1–162. Kyōto: Nakanishi Printing.
- Izuyama, Atsuko. 2012. Yonaguni. In Nicolas Tranter (ed.), *The languages of Japan and Korea* (Routledge Language Family Series), 412–457. London: Routledge.
- Jacobs, Neil G. 2005. *Yiddish. A linguistic introduction*. Cambridge: Cambridge University Press.
- Jacobs, Neil G., Ellen F. Prince & Johan van der Auwera. 1994. Yiddish. In Ekkehard König & Johan van der Auwera (eds.), *The Germanic languages* (Routledge Language Family Series), 388–419. London: Routledge.
- Jacobson, Steven. 1979. *A grammatical sketch of Siberian Yupik Eskimo*. Fairbanks: Alaska Native Language Center.
- Jacobson, Steven. 2001. *A practical grammar of the St. Lawrence Island / Siberian Yupik Eskimo language*. 2nd edn. Fairbanks: Alaska Native Language Center.
- Jacques, Guillaume. 2004. *Phonologie et morphologie du Japhug (rGyalrong)*. Paris: Université Paris VII. (Doctoral dissertation.)
- Jacques, Guillaume. 2011. The structure of the Tangut verb. *Journal of Chinese Linguistics* 39(2). 419–441.
- Jacques, Guillaume. 2014. Cone. In Jackson T.-S. Sun (ed.), *Phonological profiles of little-studied Tibetic varieties*, 265–371. Taipei: Academia Sinica.
- Janhunen, Juha. 1977. *Samojedischer Wortschatz. Gemeinsamojedische Etymologien* (Castrenianumin toimitteia 17). Helsinki: University of Helsinki.
- Janhunen, Juha. 1991. *Material on Manchurian Khamnigan Evenki* (Castrenianumin toimitteia 222). Helsinki: University of Helsinki.
- Janhunen, Juha. 1996. *Manchuria. An ethnic history* (Mémoires de la Société Finno-Ougrienne). Helsinki: The Finno-Ugrian Society.
- Janhunen, Juha. 1997. The languages of Manchuria in today's China. *Senri Ethnological Studies* 44. 123–146.
- Janhunen, Juha. 1998. Samoyedic. In Daniel Abondolo (ed.), *The Uralic languages* (Routledge Language Family Series), 457–479. London: Routledge.
- Janhunen, Juha. 2003a. A framework for the study of Japanese language origins. In Osada Toshiki, Alexander Vovin & Kerri Russell (eds.), *Perspectives on the origins of the Japanese language*, 477–490. Kyōto: International Research Center for Japanese Studies.
- Janhunen, Juha. 2003b. Khamnigan Mongol. In Juha Janhunen (ed.), *The Mongolic languages* (Routledge Language Family Series), 83–101. London: Routledge.
- Janhunen, Juha. 2003c. Para-Mongolic. In Juha Janhunen (ed.), *The Mongolic languages* (Routledge Language Family Series), 391–402. London: Routledge.

- Janhunen, Juha. 2003d. Proto-Mongolic. In Juha Janhunen (ed.), *The Mongolic languages* (Routledge Language Family Series), 1–29. London: Routledge.
- Janhunen, Juha (ed.). 2003e. *The Mongolic Languages* (Routledge Language Family Series). London: Routledge.
- Janhunen, Juha. 2003f. Written Mongol. In Juha Janhunen (ed.), *The Mongolic languages* (Routledge Language Family Series), 30–56. London: Routledge.
- Janhunen, Juha. 2005. Tungusic. An endangered language family in Northeast Asia. *International Journal of the Sociology of Language* 173. 37–54.
- Janhunen, Juha. 2006. The Mongolic languages. In Keith Brown (ed.), *Encyclopedia of language & linguistics*, 2nd edn., 231–234. Oxford: Elsevier.
- Janhunen, Juha. 2007a. Typological interaction in the Qinghai linguistic complex. *Studia Orientalia* 101. 85–102.
- Janhunen, Juha. 2007b. Typological expansion in the Ural-Altai belt. *Incontri linguistici* 30. 71–83.
- Janhunen, Juha. 2009. Proto-Uralic: What, where, and when? *Mémoires de la Société Finno-Ougrienne* 258. 57–78.
- Janhunen, Juha. 2010. Reconstructing the language map of prehistorical Northeast Asia. *Studia Orientalia* 108. 281–303.
- Janhunen, Juha. 2012a. Khitan. Understanding the language behind the scripts. *SCRIPTA* 4. 107–132.
- Janhunen, Juha. 2012b. *Mongolian* (London Oriental and African Language Library 19). Amsterdam: Benjamins.
- Janhunen, Juha. 2012c. On the hierarchy of structural convergence in the Amdo Sprachbund. In Pirkko Suihkonen, Bernard Comrie & Valery Solovyev (eds.), *Argument structure and grammatical relations: A crosslinguistic typology* (Studies in Language Companion Series 126), 177–190. Amsterdam: Benjamins.
- Janhunen, Juha. 2012d. The expansion of Tungusic as an ethnic and linguistic process. In Andrej Malchokov & Lindsay J. Whaley (eds.), *Recent advances in Tungusic linguistics* (Turcologica 89), 5–16. Wiesbaden: Harrassowitz.
- Janhunen, Juha. 2013. Personal pronouns in core Altaic. In Martine Robbeets & Hubert Cuyckens (eds.), *Shared grammaticalization. With special focus on the Transeurasian languages* (Studies in Language Companion Series 132), 211–226. Amsterdam: Benjamins.
- Janhunen, Juha, Marja Peltomaa, Erika Sandman & Xiawu Dongzhu. 2008. *Wutun* (Languages of the World/Materials 466). Munich: Lincom Europa.
- Järvilehto, Timo. 1998. The theory of the organism-environment system I. Description of the theory. *Integrative Physiological and Behavioral Science* 33(4). 321–334.
- Järvilehto, Timo. 2009. The theory of the organism-environment system as a basis of experimental work in psychology. *Ecological Psychology* 21(2). 112–120.
- Jedig, Hugo H. 2014. *Lepel, Laumtje, Lostichkeit. Gesammelte Beiträge zu deutschen Mundarten in der Sowjetunion*. Mannheim: Institut für deutsche Sprache.
- Jeon, Hae-Sung. 2015. Prosody. In Lucien Brown & Yeon Jaehoon (eds.), *The handbook of Korean linguistics*, 41–58. Malden: Wiley Blackwell.

## References

- Jiang Li. 2015. *A grammar of Guiqióng. A language of Sichuan* (Brill's Tibetan Studies Library Sub Series. Languages of the Greater Himalayan Region 15). Leiden: Brill.
- Jinam, Timothy, Nishida Nao, Hirai Momoki, Kawamura Shoji, Oota Hiroki, Umetsu Kazuo, Kimura Ryosuke, Ohashi Jun, Tajima Atsushi, Yamamoto Toshimichi, Tanabe Hideyuki, Mano Shuhei, Suto Yumiko, Kaname Tadashi, Naritomi Kenji, Yanagi Kumiko, Niikawa Norio, Omoto Keiichi, Tokunaga Katsushi & Saitou Naruya. 2012. The history of human populations in the Japanese archipelago inferred from genome-wide SNP data with a special reference to the Ainu and the Ryukyuan populations. *Journal of Human Genetics* 57. 787–795.
- Johanson, Lars. 1998. The history of Turkic. In Lars Johanson & Éva Á. Csató (eds.), *The Turkic languages* (Routledge Language Family Series), 81–125. London: Routledge.
- Johanson, Lars. 2006a. Turkic languages. In Keith Brown (ed.), *Encyclopedia of language & linguistics*, 2nd edn., 161–164. Oxford: Elsevier.
- Johanson, Lars. 2006b. Uzbek. In Keith Brown (ed.), *Encyclopedia of language & linguistics*, 2nd edn., 287–290. Oxford: Elsevier.
- Johanson, Lars. 2010. Turkic language contacts. In Raymond Hickey (ed.), *The handbook of language contact*, 652–672. Chichester: Wiley-Blackwell.
- Joki, A. J. 1944. *Kai Donners kamassisches Wörterbuch nebst Sprachproben und Hauptzügen der Grammatik*. Helsinki: Suomalais-Ugrilainen Seura.
- Jones, Eppie R., Gloria Gonzales-Fortes, Sarah Connell, Veronika Siska, Anders Eriksson, Rui Martiniano, Russell L. McLaughlin, Marcos Gallego Llorente, Lara M. Cassidy, Cristina Gamba, Tengiz Meshveliani, Ofer Bar-Yosef, Werner Müller, Anna Belfer-Cohen, Zinovi Matskevich, Nino Jakeli, Thomas F. G. Higham, Mathias Currat, David Lordkipanidze, Michael Hofreiter, Andrea Manica, Ron Pinhasi & Daniel G. Bradley. 2015. Upper palaeolithic genomes reveal deep roots of modern Eurasians. *Nature Communications* 6(8912). 1–8.
- Kałużyński, Stanisław. 1971. Solonisches Wörterverzeichnis. *Rocznik Orientalistyczny* 34(2). 15–56.
- Kałużyński, Stanisław. 1977. Sibirische Sprachproben. III. Sprachbemerkungen. *Rocznik Orientalistyczny* 39(1). 35–53.
- Kämpfe, Hans-Rainer & Alexander P. Volodin. 1995. *Abriß der tschuktschischen Grammatik auf Basis der Schriftsprache* (Tunguso Sibirica 1). Wiesbaden: Harrassowitz.
- Kane, Daniel. 1989. *The Sino-Turkic vocabulary of the Bureau of Interpreters*. Bloomington: Indiana University Press.
- Kang, Min Jeong, Ming Hsu, Ian M. Krajbich, George Loewenstein, Samuel M. McClure, Wang Joseph Tao-yi & Colin F. Camerer. 2009. The wick in the candle of learning. Epistemic curiosity activates reward circuitry and enhances memory. *Psychological Science* 20(8). 963–973.
- Kara, Dávid Somfai. 2003. *Kyrgyz* (Languages of the World/Materials 423). Munich: Lincom Europa.
- Karlsson, Anastasia Mukhanova. 2003. Tonal gestures in Mongolian interrogatives. *Phonum* 9. 189–192.
- Katz, Dovid. 1987. *Grammar of the Yiddish language*. London: Duckworth.

- Kazama, Shinjirō. 2002a. *Negidarugo tekisuto to bunpō gaisetsu* (Publications on Tungus Languages and Cultures 19). Kyōto: Nakanishi.
- Kazama, Shinjirō. 2002b. *Urucha kōshō bungei genbun-shū* (Publications on Tungus Languages and Cultures 20). Kyōto: Nakanishi.
- Kazama, Shinjirō. 2003. *Basic vocabulary (A) of Tungusic languages* (Endangered Languages of the Pacific Rim A2-037). Kyōto: Nakanishi.
- Kazama, Shinjirō. 2007. Rhetorical questions with interrogative markers in Nanai. In Kawaguchi Yuji, Takagaki Toshihiro, Tomimori Nobuo & Tsuruga Yoichiro (eds.), *Corpus-based perspectives in linguistics* (Usage-Based Linguistic Informatics), 319–335. Amsterdam: Benjamins.
- Khabtagaeva, Bayarma. 2017. *The Ewenki dialects of Buryatia and their relationship to Khamnigan Mongol* (Tunguso Sibirica 41). Wiesbaden: Harrassowitz.
- Khalilova, Zaira. 2009. *A grammar of Khwarshi* (LOT 234). Utrecht: LOT.
- Khasanova, Marina & Alexander M. Pevnov. 2003. *Myths and tales of the Negidals* (Lincom Studies in Asian Linguistics 21). Kyōto: Nakanishi.
- Kiaer, Jieun. 2014. *Jeju language and tales from the edge of the Korean peninsula* (Lincom Studies in Asian Linguistics 83). Munich: Lincom.
- Kiefer, Ferenc. 1980. Yes-no questions as WH-questions. In John R. Searle, Ferenc Kiefer & Manfred Bierwisch (eds.), *Speech act theory and pragmatics* (Studies in Linguistics and Philosophy 10), 79–119. Dordrecht: Reidel.
- Kim, Deborah. 2014. *Subordination in Sarikoli*. Grand Forks: University of North Dakota. (Master's thesis.)
- Kim, Juwon, Dongho Ko, Chaoke D. O., Han Youfeng, Piao Lianyu & Boris V. Boldyrev. 2008. *Materials of spoken Manchu*. Seoul: Seoul National University Press.
- Kim, Ronald I. 2012. *Introduction to Tocharian*. Prague: Univerzita Karlova. (Unpublished manuscript.)
- Kim, Stephen S. 2003. Santa. In Juha Janhunen (ed.), *The Mongolic languages* (Routledge Language Family Series), 346–363. London: Routledge.
- Kimball, Geoffrey D. 1991. *Koasati grammar* (Studies in the Anthropology of North American Indians). Lincoln: University of Nebraska Press.
- Kim-Renaud, Young-Key. 2012. Modern Korean. In Nicolas Tranter (ed.), *The languages of Japan and Korea* (Routledge Language Family Series), 123–167. London: Routledge.
- King, Ross J. 1987. An introduction to Soviet Korean. *Language Research* 23(2). 233–271.
- King, Ross J. 2006a. Dialectal variation in Korean. In Ho-Min Sohn (ed.), *Korean language in culture and society*, 264–280. Honolulu: University of Hawai'i Press.
- King, Ross J. 2006b. Korean dialects in the former USSR: Reflections on the current state of research. *Pangenhak* 3. 127–153.
- Kirchner, Mark. 1998a. Kazakh and Karakalpak. In Lars Johanson & Éva Á. Csátó (eds.), *The Turkic languages* (Routledge Language Family Series), 318–332. London: Routledge.
- Kirchner, Mark. 1998b. Kirghiz. In Lars Johanson & Éva Á. Csátó (eds.), *The Turkic languages* (Routledge Language Family Series), 344–356. London: Routledge.

## References

- Kiyose, Gisaburo N. 1977. *A study of the Jurchen language and script: Reconstruction and decipherment*. Kyōto: Shōwado.
- Kiyose, Gisaburo N. 2000. Genealogical relationship of Jurchen dialects and literary Manchu. *Central Asiatic Journal* 44(2). 177–189.
- Klamer, Marian. 2010. *A grammar of Teiwa* (Mouton Grammar Library 49). Berlin: De Gruyter Mouton.
- Ko, Dongho & Gyudong Yurn. 2011. *A description of Najkhin Nanai* (Altaic Languages Series 7). Seoul: Seoul National University Press.
- Kobayashi, Masato & Bablu Tirkey. 2017. *The Kurux language: Grammar, texts, and lexicon*. Vol. 08 (Brill's Studies in South and Southwest Asian Languages). Leiden: Brill.
- Köhler, Bernhard. 2013. Interrogative zero-marking in some Ometo languages. In Marie-Claude Simeone-Senelle & Martine Vanhove (eds.), *Proceedings of the 5th International Conference on Cushitic and Omotic Languages. Paris, 16-18 April 2008*, 2nd edn., 135–155. Cologne: Rüdiger Köppe Verlag.
- Köhler, Bernhard. 2016. *Form und Funktion von Fragesätzen in afrikanischen Sprachen* (Schriften zur Afrikanistik 25). Frankfurt am Main: Peter Lang.
- Kokuritsu Kokugo Kenkyūjo. 1950. *Hachijō-jima no gengo chōsa*. Tōkyō: Kokushokankōkai.
- Koloskova, Yulia & Otori Toshio. 2008. Pragmatic factors in the development of a switch-adjective language. A case study of the Miyako-Hirara dialect of Ryukyuan. *Studies in Language* 32(3). 610–636.
- König, Ekkehard & Peter Siemund. 2007. Speech act distinctions in grammar. In Timothy Shopen (ed.), *Language typology and syntactic description, vol. 1: Complex constructions*, 2nd edn., 276–324. Cambridge: Cambridge University Press.
- Kortt, I. R. & Ju. B. Simčenko. 1985. *Wörterverzeichnis der nganasanischen Sprache* (Systemata mundi/Materialien 1). Berlin: Systemata Mundi.
- Kotorova, Elizaveta & Andrey Nefedov. 2015. *Comprehensive dictionary of Ket* (Languages of the World/Dictionaries 57). Munich: Lincom. 2 vols.
- Kraaijenbrink, Thirsa, Emma J. Parkin, Denise R. Carvalho-Silva, George L. van Driem, Guido Barbujani, Chris Tyler-Smith, Mark A. Jobling & Peter de Knijff. 2009. Genetic and linguistic borders in the Himalayan region. In Francesco d'Errico & Jean-Marie Hombert (eds.), *Becoming eloquent: Advances in the emergence of language, human cognition, and modern cultures*, 181–201. Amsterdam: Benjamins.
- Krasovitsky, Alexander. 2004. *Russian in East Siberia and Kamchatka*. <http://www.ruhr-uni-bochum.de/lilab/krasovitsky/studies/dialects.htm>. (Accessed 2017-07-12.)
- Kratochvíl, František. 2007. *A grammar of Abui. A Papuan language of Alor* (LOT 157). Utrecht: LOT.
- Krause, Wolfgang & Werner Thomas. 1960. *Tocharisches Elementarbuch, vol. 1: Grammatik*. Heidelberg: Winter.
- Krauss, Michael E. Forthcoming. *Eyak grammar*. <http://www.uaf.edu/anla/collections/search/resultDetail.xml?id=EY961K2015>. (Accessed 2016-06-01.)
- Kroonen, Guus. 2013. *Etymological dictionary of Proto-Germanic* (Leiden Indo-European Etymological Dictionary Series). Leiden: Brill.

- Künnap, Ago. 1999a. *Enets* (Languages of the World/Materials 186). Munich: Lincom Europa.
- Künnap, Ago. 1999b. *Kamass* (Languages of the World/Materials 185). Munich: Lincom Europa.
- Kupchik, John E. 2011. *A grammar of the eastern old Japanese dialect*. Honolulu: University of Hawai'i. (Doctoral dissertation.)
- Kurpaska, Maria. 2010. *Chinese language(s) a look through the prism of the great dictionary of modern Chinese dialects* (Trends in Linguistics. Studies and Monographs 215). Berlin: De Gruyter Mouton.
- Ladstätter, Otto & Andreas Tietze. 1994. *Die Abdal (Äynu) in Xinjiang* (Sitzungsberichte, Österreichische Akademie der Wissenschaften Wien. Philosophisch-Historische Klasse 604). Vienna: Verlag der Österreichischen Akademie der Wissenschaften.
- Landmann, Angelika. 2009. *Türkisch. Kurzgrammatik*. Wiesbaden: Harrassowitz.
- Landmann, Angelika. 2010. *Usbekisch. Kurzgrammatik*. Wiesbaden: Harrassowitz.
- Landmann, Angelika. 2014. *Tschuwaschisch. Kurzgrammatik*. Wiesbaden: Harrassowitz.
- Langacker, Ronald W. 2008. *Cognitive grammar. A basic introduction*. Oxford: Oxford University Press.
- LaPolla, Randy. 2013. Eastern Asia: Sino-Tibetan linguistic history. In Peter Bellwood (ed.), *The global prehistory of human migration* (Routledge Language Family Series), 204–208. Malden: Wiley Blackwell.
- LaPolla, Randy. 2016. Once again on methodology and argumentation in linguistics: Problems with the arguments for recasting Sino-Tibetan as “Trans-Himalayan”. *Linguistics of the Tibeto-Burman Area* 39(2). 282–297.
- LaPolla, Randy & Huang Chenglong. 2003. *A grammar of Qiang with annotated texts and glossary* (Mouton Grammar Library 31). Berlin: Mouton de Gruyter.
- Lawrence, Wayne P. 2012. Southern Ryukyuan. In Nicolas Tranter (ed.), *The languages of Japan and Korea* (Routledge Language Family Series), 381–411. London: Routledge.
- Lbova, Ludmila. 2014. The upper paleolithic of Northeast Asia. In Colin Renfrew & Paul Bahn (eds.), *The Cambridge world prehistory, 707–723*. Cambridge: Cambridge University Press. 3 vols.
- Lee-Smith, Mei W. 1996a. The Ejnu language. In Stephen A. Wurm, Peter Mühlhäusler & Darrell T. Tyron (eds.), *Atlas of languages of intercultural communication in the Pacific, Asia, and the Americas, vol. 2(1): Texts* (Trends in Linguistics. Documentation 13), 851–863. Berlin: Walter de Gruyter.
- Lee-Smith, Mei W. 1996b. The Hezhou language. In Stephen A. Wurm, Peter Mühlhäusler & Darrell T. Tyron (eds.), *Atlas of languages of intercultural communication in the Pacific, Asia, and the Americas, vol. 2(1): Texts* (Trends in Linguistics. Documentation 13), 865–873. Berlin: Walter de Gruyter.
- Lee-Smith, Mei W. 1996c. The Tangwang language. In Stephen A. Wurm, Peter Mühlhäusler & Darrell T. Tyron (eds.), *Atlas of languages of intercultural communication in the Pacific, Asia, and the Americas, vol. 2(1): Texts* (Trends in Linguistics. Documentation 13), 875–882. Berlin: Walter de Gruyter.

## References

- Lee-Smith, Mei W. & Stephen A. Wurm. 1996. The Wutun language. In Stephen A. Wurm, Peter Mühlhäusler & Darrell T. Tyron (eds.), *Atlas of languages of intercultural communication in the Pacific, Asia, and the Americas, vol. 2(1): Texts* (Trends in Linguistics. Documentation 13), 883–897. Berlin: Walter de Gruyter.
- Lehmann, Christian. 1974. *Proto-Indo-European syntax*. Austin: University of Texas Press.
- Levinson, Stephen C. 2006. On the human “interaction engine”. In Nicholas J. Enfield & Stephen C. Levinson (eds.), *Roots of human sociality. Culture, cognition and interaction* (Wenner-Gren International Symposium Series 1), 39–69. Oxford: Berg.
- Levinson, Stephen C. 2010. Questions and responses in Yéli Dnye, the Papuan language of Rossel Island. *Journal of Pragmatics* 42. 2741–2755.
- Levinson, Stephen C. 2012a. Interrogative intimations: On a possible social economics of interrogatives. In Jan P. de Ruiter (ed.), *Questions. Formal, functional and interactional perspectives*, 11–32. Cambridge: Cambridge University Press.
- Levinson, Stephen C. 2012b. The original sin of cognitive science. *Topics in Cognitive Science* 4. 396–403.
- Lewin, Kurt. 1936. *Principles of topological psychology*. New York: McGraw-Hill.
- Li Fengxiang. 2005. Contact, attrition, and structural shift: Evidence from Oroqen. *International Journal of the Sociology of Language* 173. 55–74.
- Li Linjing. 2011. Hojen-go no kaiwa tekisuto (1). *Hoppō gengo kenkyū* 1. 197–216.
- Li Yong-Sōng. 2011. *A study of Dolgan* (Altaic Language Studies 5). Seoul: Seoul National University Press.
- Li Yong-Sōng, Lee Ho-Yong, Choi Hyong-Won, Kim Geon Sook, Lee Dong Eun & Mehmet Ölmez. 2008. *A study of the Middle Chulym dialect of the Chulym language* (Altaic Language Studies 3). Seoul: Seoul National University Press.
- Li, Charles N. & Sandra A. Thompson. 1981. *Mandarin Chinese*. Berkeley: University of California Press.
- Lichtenberk, Frantisek. 2007. A typologically unusual interrogative word in Toqabaqita and other Oceanic languages. *Oceanic Linguistics* 46(2). 603–612.
- Lie, Hiu. 1978. Solonisches Material aus dem Huin-gol. In Gerhard Doerfer & Michael Weiers (eds.), *Tungusica, vol. 1*, 126–78. Wiesbaden: Otto Harrassowitz.
- Liljegren, Henrik. 2016. *A grammar of Palula* (Studies in Diversity Linguistics 8). Berlin: Language Science Press.
- Lin Lianyun. 1985. *Salayu jianzhi*. Peking: Minzu chubanshe.
- Lindström, Eva. 1995. Animacy in interrogative pronouns. In Inger Moen, Hanne G. Simonsen & Helge Lødrup (eds.), *Papers from the XVth Scandinavian Conference of Linguistics*, 307–315. Oslo: University of Oslo.
- Ling Chunsheng. 1934. *Songhuajiang xiayou de hezhezu*. Nanjing: Guoli zhongyang yanjiuyuan lishi yuyan yanjiusuo.
- Liu Liji. 1989. *Huizu wulumuqi yuyan zhi*. Urumqi: Xinjiang daxue chubanshe.
- Liu Zhaoxiong. 1981. *Dongxiangyu jianzhi*. Peking: Minzu chubanshe.
- Llamas, Bastien, Lars Fehren-Schmitz, Guido Valverde, Julien Soubrier, Swapan Mallick, Nadin Rohland, Susanne Nordenfelt, Cristina Valdiosera, Stephen M. Richards, Adam Rohrlach, Maria Inés Barreto Romero, Isabel Flores Espinoza, Elsa Tomasto Cagigao,

- Lucía Watson Jiménez, Krzysztof Makowski, Ilán Santiago Leboreiro Reyna, Josefina Mansilla Lory, Julio Alejandro Ballivián Torrez, Mario A. Rivera, Richard L. Burger, Maria Constanza Ceruti, Johan Reinhard, R. Spencer Wells, Gustavo Politis, Calogero M. Santoro, Vivien G. Standen, Colin I. Smith, Ho Simon Y. W. David Reich, Alan Cooper & Wolfgang Haak. 2016. Ancient mitochondrial DNA provides high-resolution time scale of the peopling of the Americas. *Science Advances* 2(4). 1–10.
- Loewenstein, George. 1994. The psychology of curiosity: A review and reinterpretation. *Psychological Bulletin* 116(1). 75–98.
- Lopatin, Ivan A. 1957. Material on the Orochee language. *Micro-Bibliotheca Anthropos* 26. 1–110.
- Luo Tianhua. 2013. *Interrogative strategies. An areal typology of the languages of China*. Konstanz: Universität Konstanz. (Doctoral dissertation.)
- Ma Guoliang & Liu Zhaoxiong. 1986. Dongxiangyu yanjiu. *Xibei minzu yanjiu* 00. 167–184, 232.
- Ma Quanlin, Ma Wanxiang & Ma Zhicheng. 1993. Salar language materials. *Sino-Platonic Papers* 43. 1–72.
- Mackenzie, J. Lachlan. 2009. Content interrogatives in a sample of 50 languages. *Lingua* 119. 1131–1163.
- Maddieson, Ian. 2013. Front rounded vowels. In Matthew S. Dryer & Martin Haspelmath (eds.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://wals.info/chapter/11>. (Accessed 2015-03-02.)
- Majewicz, Alfred F. (ed.). 1998. *The collected works of Bronisław Piłsudski, vol. 2: Materials for the study of the Ainu language and folklore, vol. 1*. Berlin: Mouton de Gruyter.
- Majewicz, Alfred F. (ed.). 2011. *The collected works of Bronisław Piłsudski, vol. 4: Materials for the study of Tungusic languages and folklore*. Berlin: Mouton de Gruyter.
- Makelaike Yumai'erbai. 1986. Shilun xinjiang keerkeziyu fangyan. *Yuyan yu fanyi* 2. 21–24, 35.
- Malchukov, Andrej L. 1995. *Even* (Languages of the World/Materials 12). Munich: Lincom Europa.
- Malchukov, Andrej L. 2001. Imperative constructions in Even. In Victor S. Xrakovskij (ed.), *Typology of imperative constructions* (Lincom Studies in Theoretical Linguistics 9), 159–180. Munich: Lincom Europa.
- Malchukov, Andrej L. 2008. *Sintaksis evenskogo jazyka*. Sankt-Peterburg: Nauka.
- Malchukov, Andrej L. & Igor Nedjalkov. 2010. Ditransitive constructions in Tungusic languages. In Andrej Malchukov, Martin Haspelmath & Bernard Comrie (eds.), *Studies in ditransitive constructions. A comparative handbook*, 316–351. Berlin: Mouton de Gruyter.
- Mallory, James P. 2010. Bronze age languages of the Tarim Basin. *Expedition* 52(3). 44–53.
- Mallory, James P. & Douglas Q. Adams. 2006. *The Oxford introduction to Proto-Indo-European and the Proto-Indo-European world* (Oxford Linguistics). Oxford: Oxford University Press.

## References

- Manzelli, Gianguido. 2015. Mutual influences in negative patterns between Finno-Ugric and Turkic languages in the Volga-Kama area. In Matti Miestamo, Anne Tamm & Beáta Wagner-Nagy (eds.), *Negation in Uralic languages* (Typological Studies in Language 108), 633–652. Amsterdam: Benjamins.
- Martin, Samuel E. 1961. *Dagur Mongolian grammar, texts, and lexicon. Based on the speech of Peter Onon* (Indiana University Uralic and Altaic Series 4). Bloomington: Indiana University.
- Martin, Samuel E. 1970. Shodon. A dialect of the Northern Ryukyus. *Journal of the American Oriental Society* 90(1). 97–139.
- Maslova, Elena. 1997. Yukagir focus in a typological perspective. *Journal of Pragmatics* 27. 457–475.
- Maslova, Elena. 2001. *Yukaghir texts* (Tunguso-Sibirica 7). Wiesbaden: Harrassowitz.
- Maslova, Elena. 2003a. *A grammar of Kolyma Yukaghir* (Mouton Grammar Library 27). Berlin: Mouton de Gruyter.
- Maslova, Elena. 2003b. *Tundra Yukaghir* (Languages of the World/Materials 372). Munich: Lincom Europa.
- Matayoshi, Satomi. 2010a. Hateruma (Yaeyama Ryukyuan). In Thomas Pellard & Michinori Shimoji (eds.), *An introduction to Ryukyuan languages*, 189–227. Tōkyō: Research Institute for Languages, Cultures of Asia & Africa.
- Matayoshi, Satomi. 2010b. Tsuken (Okinawan). In Thomas Pellard & Michinori Shimoji (eds.), *An introduction to Ryukyuan languages*, 89–111. Tōkyō: Research Institute for Languages, Cultures of Asia & Africa.
- Matić, Dejan. 2014. Questions and syntactic islands in Tundra Yukaghir. In Rik van Gijn, Jeremy Hammond, Dejan Matić, Saskia van Putten & Ana V. Galucio (eds.), *Information structure and reference tracking in complex sentences* (Typological Studies in Language 105), 127–162. Amsterdam: Benjamins.
- Matić, Dejan. 2016. Tag questions and focus markers: Evidence from the Tompo dialect of Even. In M. M. Jocelyne Fernandez-Vest & Robert D. Van Valin (eds.), *Information structure and spoken language in a cross-linguistic perspective* (Trends in Linguistics. Studies and Monographs 283), 167–189. Berlin: Mouton de Gruyter.
- Matić, Dejan & Brigitte Pakendorf. 2013. Non-canonical SAY in Siberia. Areal and genealogical patterns. *Studies in Language* 37(2). 356–412.
- Matsumori, Akiko & Onishi Takuichiro. 2012. Japanese dialects: Focusing on Tsuruoka and Ei. In Nicolas Tranter (ed.), *The languages of Japan and Korea* (Routledge Language Family Series), 313–348. London: Routledge.
- Matthews, Stephen. 2010. Language contact and Chinese. In Raymond Hickey (ed.), *The handbook of language contact* (Routledge Language Family Series), 757–769. Chichester: Wiley-Blackwell.
- Mattissen, Johanna. 2003. *Dependent-head synthesis in Nivkh. A contribution to a typology of polysynthesis* (Typological Studies in Language 57). Amsterdam: Benjamins.
- Mauri, Caterina. 2008. The irrealty of alternatives. Towards a typology of disjunction. *Studies in Language* 32(1). 22–55.

- Mawkanuli, Talant. 2005. *Jungar Tuvan texts* (Indiana University Uralic and Altaic Series 170). Bloomington: Indiana University.
- McWhorter, John. 2007. *Language interrupted. Signs of non-native acquisition in standard language grammars*. Oxford: Oxford University Press.
- Menovshchikov, G. A. 1975. *Jazyk naukanskix eskimosov*. Leningrad: Nauka.
- Meyer, Michel. 1980. Science as a questioning process: A prospect for a new type of rationality. *Revue Internationale de Philosophie* 34(131/132). 49–89.
- Mi Haili. 1997. *Weiwueryu kashihua yanjiu*. Peking: Zhongyang minzudaxue chubanshe.
- Miestamo, Matti. 2008. Grammatical complexity in a cross-linguistic perspective. In Matti Miestamo, Kaius Sinnemäki & Fred Karlsson (eds.), *Language complexity: Typology, contact, change* (Studies in Language Companion Series 94), 23–41. Amsterdam: Benjamins. 5 vols.
- Miestamo, Matti. 2011. Polar interrogatives in Uralic languages. A typological perspective. *Linguistica Uralica* 47(1). 1–21.
- Mikola, Tibor. 2004. *Studien zur Geschichte der samojedischen Sprachen* (Studia Uralo-Altaica 45). Szeged: SzTE Finnisch-Ugrisches Institut.
- Mithun, Marianne. 2012. Tags: Cross-linguistic diversity and commonality. *Journal of Pragmatics* 44. 2165–2182.
- Miyake, Marc. 2017. Khitan language. In Rint Sybesma (ed.), *Encyclopedia of Chinese language and linguistics*, 492–495. Leiden: Brill. 5 vols.
- Miyaoka, Osahito. 2012. *A grammar of Central Alaskan Yupik (CAY)* (Mouton Grammar Library 58). Berlin: De Gruyter Mouton.
- Miyara, Shinsho. 2015. Shuri Okinawan grammar. In Heinrich Patrick, Shinsho Miyara & Michinori Shimoji (eds.), *Handbook of the Ryukyuan languages. History, structure, and use* (Handbooks of Japanese Language and Linguistics 11), 379–404. Berlin: De Gruyter Mouton.
- Moravcsik, Edith A. 1971. Some cross-linguistic generalizations about yes-no questions and their answers. *Working Papers on Language Universals* 7. 45–193.
- Moreno-Mayar, J. Víctor, Ben A. Potter, Lasse Vinner, Matthias Steinrücken, Simon Rasmussen, Jonathan Terhorst, John A. Kamm, Anders Albrechtsen, Anna-Sapfo Malaspina, Martin Sikora, Joshua D. Reuther, Joel D. Irish, Ripan S. Malhi, Ludovic Orlando, Yun S. Song, Rasmus Nielsen, David J. Meltzer & Eske Willerslev. 2018. Terminal Pleistocene Alaskan genome reveals first founding population of Native Americans. *Nature* 553. 1–6.
- Mostaert, Antoine. 1937. *Textes oraux ordos*. Beijing: The Catholic University of Peking.
- Mu Yejun. 1985. Alechuka manyu yuyin jianlun. *Manyu yanjiu* 1. 5–15.
- Mu Yejun. 1986. Alechuka manyu de shuci yu gezhuci. *Manyu yanjiu* 2. 2–17.
- Mu Yejun. 1987. Balayu. *Manyu yanjiu* 2. 2–31, 128.
- Mu Yejun. 1988a. Alechukayu yuanyin fasheng de yinbian tedian. *Manyu yanjiu* 7. 1–24.
- Mu Yejun. 1988b. Lun balayu de yuyin bianhua. *Manyu yanjiu* 6. 1–26.
- Muhamedowa, Raihan. 2016. *Kazakh. A comprehensive grammar* (Routledge Comprehensive Grammars). London: Routledge.

## References

- Mühlhäusler, Peter. 1992. Preserving languages or language ecologies? A top-down approach to language survival. *Oceanic Linguistics* 31(2). 163–180.
- Mus, Nikolett. 2009. The question-words in Tundra Nenets. I. The nominal question words. *Ural-Altäische Forschungen* 1. 43–48.
- Mus, Nikolett. 2013. The typology of Northern Samoyed question words. *Finnisch-Ugrische Mitteilungen* 37. 175–193.
- Mus, Nikolett. 2015a. *Interrogative words and content questions in Tundra Nenets*. Szeged: University of Szeged. (Doctoral dissertation).
- Mus, Nikolett. 2015b. Negation in Tundra Nenets. In Matti Miestamo, Anne Tamm & Beáta Wagner-Nagy (eds.), *Negation in Uralic languages* (Typological Studies in Language 108), 75–101. Amsterdam: Benjamins.
- Mushin, Ilana. 1995. Epistememes in Australian languages. *Australian Journal of Linguistics* 15. 1–31.
- Muysken, Pieter & Norval Smith. 1990. Question words in pidgin and creole languages. *Linguistics* 28. 883–903.
- Nagano, Yasuhiko. 1980. *Amdo Sherpa dialect. A material for Tibetan dialectology* (Monumenta Serindica 7). Tōkyō: Institute for the Study of Languages, Cultures of Asia & Africa.
- Nagano-Madsen, Yasuko. 2015. Intonation in Okinawan. In Heinrich Patrick, Shinsho Miyara & Michinori Shimoji (eds.), *Handbook of the Ryukyuan languages: History, structure, and use* (Handbooks of Japanese Language and Linguistics 11), 199–226. Berlin: De Gruyter Mouton.
- Nagasaki, Iku. 2011. Kolyma Yukaghir. In Yasuhiro Yamakoshi (ed.), *Grammatical sketches from the field*, 213–256. Tōkyō: Research Institute for Languages, Cultures of Asia & Africa.
- Nagayama, Yukari. 2011. Alutor. In Yasuhiro Yamakoshi (ed.), *Grammatical sketches from the field*, 257–302. Tōkyō: Research Institute for Languages, Cultures of Asia & Africa.
- Nagayama, Yukari. 2016. Nominalization in Alutor. *Studia Orientalia* 117. 119–136.
- Nakagawa, Hiroshi & Osami Okuda. 2007. Nivkh and Ainu. In Miyaoka Osahito, Sakiyama Osamu & Michael E. Krauss (eds.), *The vanishing languages of the Pacific Rim* (Oxford Linguistics), 371–386. Oxford: Oxford University Press.
- Nakanome, Akira. 1928. *Grammatik der Orokkō-Sprache*. Osaka: The Osaka Asiatic Society.
- Nam, Pung-hyun. 2012. Old Korean. In Nicolas Tranter (ed.), *The languages of Japan and Korea* (Routledge Language Family Series), 41–71. London: Routledge.
- Napoli, Mateus Froes. 2014. *Studies on the verb complex of Santa Mongolian*. Chiang Mai: Payap University. (Master's thesis.)
- Narangoa, Li & Robert Cribbs. 2014. *Historical atlas of Northeast Asia, 1590-2010. Korea, Manchuria, Mongolia, eastern Siberia*. New York: Columbia University Press.
- Nau, Nicole. 1999. Was schlägt der Kasus? Zu Paradigmen und Formengebrauch von Interrogativpronomen. *Sprachtypologie und Universalienforschung* 2. 130–150.

- NDSSLD, Neimeng dongbei shaoshuminzu shehui lishi diaochazu. 1958. *Heilongjiang-sheng fuyuanxian xiabacha hezhe minzuxiang qingkuang, hezhezudiaochoa cailiao zhi yi*. Peking: Quanguo renmin daibiao dahui minzu weiyuanhui bangongshi.
- Nedjalkov, Igor. 1997. *Evenki* (Descriptive Grammars). London: Routledge.
- Nedjalkov, Vladimir. 1994. Tense-aspect-mood forms in Chukchi. *Sprachtypologie und Universalienforschung* 47(4). 278–354.
- Nedjalkov, Vladimir P. & Galina A. Otaina. 2013. *A syntax of the Nivkh language. The Amur dialect* (Studies in Language Companion Series 139). Amsterdam: Benjamins.
- Nettle, Daniel. 1999. *Linguistic diversity* (Oxford Linguistics). Oxford: Oxford University Press.
- Nevskaja, Irina. 2000. Shor-Russian contact features. *Studies in Slavic and General Linguistics* 28. 283–298.
- Nevskaja, Irina. 2014. Locational and directional relations and tense and aspect marking in Chalkan, a South Siberian Turkic language. In Pirkko Suihkonen & Lindsay J. Whaley (eds.), *On diversity and complexity of languages spoken in Europe and North and Central Asia* (Studies in Language Companion Series 164), 67–90. Amsterdam: Benjamins.
- Nichols, Johanna. 1992. *Linguistic diversity in time and space*. Chicago: The University of Chicago Press.
- Nichols, Johanna. 1997. Modeling ancient population structures and movement in linguistics. *Annual Review of Anthropology* 26. 359–384.
- Nichols, Johanna. 2010. Macrofamilies, macroareas, and contact. In Raymond Hickey (ed.), *The handbook of language contact*, 361–379. Chichester: Wiley-Blackwell.
- Nichols, Johanna. 2015. Types of spread zones. Open and closed, horizontal and vertical. In Rik De Busser & Randy J. LaPolla (eds.), *Language structure and environment. Social, cultural, and natural factors* (Cognitive Linguistic Studies in Cultural Contexts 6), 261–286. Amsterdam: Benjamins.
- Nichols, Johanna & David A. Peterson. 2013. M-T Pronouns. In Matthew S. Dryer & Martin Haspelmath (eds.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://wals.info/chapter/136>. (Accessed 2016-04-10.)
- Nieuweboer, Rogier. 1999. *The Altai dialect of Plautdiitsch. West Siberian Menonite Low German* (Lincom Studies in Germanic Linguistics 7). Munich: Lincom Europa.
- Niinaga, Yuto. 2010. Yuwan (Amami Ryukyuan). In Thomas Pellard & Michinori Shimoji (eds.), *An introduction to Ryukyuan languages*, 35–88. Tōkyō: Research Institute for Languages, Cultures of Asia & Africa.
- Niinaga, Yuto. 2015. Amami grammar. In Heinrich Patrick, Shinsho Miyara & Michinori Shimoji (eds.), *Handbook of the Ryukyuan languages. History, structure, and use* (Handbooks of Japanese Language and Linguistics 11), 323–343. Berlin: De Gruyter Mouton.
- Nikolaeva, Irina A. 2006. *A historical dictionary of Yukaghir* (Trends in Linguistics. Documentation 25). Berlin: Mouton de Gruyter.
- Nikolaeva, Irina A. 2008. Chuvan and Omok languages. *Studies in Slavic and General Linguistics* 33. 313–336.

## References

- Nikolaeva, Irina A. 2014. *A grammar of Tundra Nenets* (Mouton Grammar Library 65). Berlin: Walter de Gruyter.
- Nikolaeva, Irina A. & Maria Tolskaya. 2001. *A grammar of Udihe* (Mouton Grammar Library 22). Berlin: Mouton de Gruyter.
- NINJAL, Kokuritsu kokugo kenkyūsho. 2015. *A topical dictionary of conversational Ainu*. <http://ainutopic.ninjal.ac.jp>. (Accessed 2016-06-11.)
- Norman, Jerry. 2013. *A comprehensive Manchu-English dictionary* (Harvard-Yenching Institute Monograph Series 85). Cambridge: Harvard University Asia Center.
- Norman, Jerry. 2014. A model for Chinese dialect evolution. In Richard VanNess Simmons & Newell Ann Van Auken (eds.), *Studies in Chinese and Sino-Tibetan linguistics: Dialect, phonology, transcription and text* (Language and Linguistics Monograph Series 53), 1–26. Taipei: Institut of Linguistics, Academia Sinica.
- Novák, Lubomír. 2014. Question of (re)classification of Eastern Iranian languages. *Linguistica Brunensia* 62(1). 77–87.
- Nugteren, Hans. 2003. Shira Yughur. In Juha Janhunen (ed.), *The Mongolic languages* (Routledge Language Family Series), 265–285. London: Routledge.
- OCLS, Okinawa Center of Language Study. 1999/2003. *Ryūkyū-go onsei dētabēsu sakusei purojekuto*. <http://ryukyū-lang.lib.u-ryukyū.ac.jp/index.html>. (Accessed 2017-07-08.)
- O'Connor, Loretta. 2014. Is the order of the interrogative word and the clause obligatorily question word - clause (i.e., are question words positioned initially)? In Pieter Muysken, Harald Hammarström, Olga Krasnoukhova, Neele Müller, Joshua Birchall, Simon van de Kerke, Loretta O'Connor, Swintha Danielsen, Rik van Gijn & George Saad (eds.), *South American Indian language structures (SAILS) online*. Leipzig: Online MPI of Evolutionary Anthropology. <http://sails.clld.org>. (Accessed 2016-04-11.)
- Odling-Smee, John, Douglas H. Erwin, Eric P. Palkovacs, Marcus W. Feldman & Kevin N. Laland. 2013. Niche construction theory: A practical guide for ecologists. *The Quarterly Review of Biology* 88(1). 3–28.
- Odling-Smee, John & Kevin N. Laland. 2009. Cultural niche construction: Evolution's cradle of language. In Rudolf Botha & Chris Knight (eds.), *The prehistory of language*, 99–121. Oxford: Oxford University Press.
- Olawsky, Knut J. 2006. *A grammar of Urarina* (Mouton Grammar Library 37). Berlin: Mouton de Gruyter.
- Ōmōryūi, Thomas O. 1988. On the formation of questions in Edo. *Journal of African Languages and Linguistics* 10. 19–31.
- Operstein, Natalie. 2015. Contact-genetic linguistics: Toward a contact-based theory of language change. *Language Sciences* 48. 1–15.
- Pakendorf, Brigitte. 2010. Contact and Siberian languages. In Raymond Hickey (ed.), *The handbook of language contact*, 714–737. Chichester: Wiley-Blackwell.
- Parker, Steve. 2006. A cross-linguistic corpus of forms meaning 'yes'. *Linguistic Discovery* 4(1). 1–34.
- Parpola, Asko. 2012. The problem of Samoyed origins in the light of archaeology: On the formation and dispersal of East Uralic (Proto-Ugro-Samoyed). *Mémoires de la Société Finno-Ougrienne* 264. 287–298.

- Pellard, Thomas. 2005. Review: Christopher I. Beckwith (2004) *Koguryo, the language of Japan's continental relatives: An introduction to the historical-comparative study of the Japanese-Koguryoic languages with a preliminary description of Archaic North-eastern Middle Chinese*. *Korean Studies* (29). 167–170.
- Pellard, Thomas. 2008. Proto-Japonic \*e and \*o in Eastern Old Japanese. *Cahiers de linguistique. Asie Orientale* 37(2). 133–158.
- Pellard, Thomas. 2009. *Ōgami. Éléments de description d'un parler du Sud des Ryūkyū*. Paris: École des Hautes Etudes en Sciences Sociales. (Doctoral dissertation.)
- Pellard, Thomas. 2010. Ōgami (Miyako Ryukyuan). In Thomas Pellard & Michinori Shimoji (eds.), *An introduction to Ryukyuan languages*, 113–166. Tōkyō: Research Institute for Languages, Cultures of Asia & Africa.
- Peng Qiu. 2015. *A preliminary investigation of Yilan Creole in Taiwan: Discussing predicate position in Yilan Creole*. Edmonton: University of Alberta. (Master's thesis.)
- Perls, Frederick S. 1973. *The gestalt approach & eye witness to therapy*. Ben Lomond: Science & Behavior Books.
- Peterson, John. 2011. *A grammar of Kharia. A south Munda language* (Brill's Studies in South and Southwest Asian Languages 1). Leiden: Brill.
- Pevnov, Alexander M. 2009. On some features of Nivkh and Uilta (in connection with prospects of Russian-Japanese collaboration). In Toshiro Tsumagari (ed.), *Linguistic world of Sakhalin: Proceedings of the symposium, September 6, 2008*, 113–125. Sapporo: Graduate School of Letters, Hokkaidō University.
- Pevnov, Alexander M. 2012. The problem of the localization of the Manchu-Tungusic homeland. In Andrej L. Malchukov & Lindsay J. Whaley (eds.), *Recent advances in Tungusic linguistics* (Turcologica 89), 17–40. Wiesbaden: Harrassowitz.
- Pevnov, Alexander M. 2016. On the specific features of Oroch as compared with the other Tungusic languages. *Studia Orientalia* 117. 47–63.
- Peyraube, Alain & Fuxiang Wu. 2005. Origin and evolution of question-words in Archaic Chinese: A cognitive approach. *Cahier Linguistique. Asie Orientale* 34(1). 3–24.
- Peyrot, Michaël. 2010. Proto-Tocharian syntax and the status of Tocharian A. *The Journal of Indo-European Studies* 38(1/2). 132–146.
- Pispane, Peter S. 2013. The Uralic-Yukaghiric connection revisited: Sound correspondences of geminate clusters. *Journal de la Société Finno-Ougrienne* 94. 165–197.
- Pitulko, Vladimir V. & Aleksey K. Kasparov. 2017. Archaeological dogs from the Early Holocene Zhokhov site in the Eastern Siberian Arctic. *Journal of Archaeological Science: Reports* 13. 491–515.
- Pitulko, Vladimir V., Alexei N. Tikhonov, Elena Y. Pavlova, Pavel A. Nikolskiy, Konstantin E. Kuper & Roman N. Polozov. 2016. Early human presence in the arctic: Evidence from 45,000-year-old mammoth remains. *Science* 351. 260–263.
- Poppe, Nicholas. 1931. *Materialy po solonskomu jazyku*. Leningrad: Nauka.
- Poppe, Nicholas. 1955. Introduction to Mongolian comparative studies. *Mémoires de la Société Finno-Ougrienne* 110. 1–300.

## References

- Poppe, Nicholas. 1963. *Tatar manual: Descriptive grammar and texts with a Tatar-English glossary* (Indiana University Uralic and Altaic series 25). Bloomington: Indiana University.
- Prins, Marielle. 2017. *A grammar of rGyalrong, Jiăomùzú (Kyom-kyo) dialects. A web of relations* (Brill's Tibetan Studies Library. Languages of the Greater Himalayan Region 16). Leiden: Brill.
- Pugh, Stefan M. & Ian Press. 1999. *Ukrainian: Comprehensive grammar* (Comprehensive Grammars). London: Routledge.
- Pulleyblank, Edwin G. 1995. *Outline of Classical Chinese grammar*. Vancouver: UBC Press.
- Ragagnin, Elisabetta. 2011. *Dukhan, a Turkic variety of northern Mongolia. Description and analysis*. Wiesbaden: Harrassowitz.
- Raghavan, Maanasa, Michael DeGiorgio, Anders Albrechtsen, Ida Moltke, Pontus Skoglund, Thorfinn S. Korneliussen, Bjarne Grønnow, Martin Appelt, Hans Christian Gulløv, T. Max Friesen, William Fitzhugh, Helena Malmström, Simon Rasmussen, Jesper Olsen, Linea Melchior, Benjamin T. Fuller, Simon M. Fahrni, Thomas W. Stafford Jr., Vaughan Grimes, M. A. Priscilla Renouf, Jerome Cybulski, Niels Lynnerup, Marta Mirazon Lahr, Kate Britton, Rick Knecht, Jette Arneborg, Mait Metspalu, Omar E. Cornejo, Anna-Sapfo Malaspinas, Wang Yong, Morten Rasmussen, Vibha Raghavan, Thomas V. O. Hansen, Elza Khusnutdinova, Tracey Pierre, Kirill Dneprovsky, Claus Andreasen, Hans Lange, M. Geoffrey Hayes, Joan Coltrain, Victor A. Spitsyn, Anders Götherström, Ludovic Orlando, Toomas Kivisild, Richard Villems, Michael H. Crawford, Finn C. Nielsen, Jørgen Dissing, Jan Heinemeier, Morten Meldgaard, Carlos D. Bustamante, Dennis H. O'Rourke, Mattias Jakobsson, M. Thomas P. Gilbert, Rasmus Nielsen & Eske Willerslev. 2014. The genetic prehistory of the new world Arctic. *Science* 345. 1–9.
- Raghavan, Maanasa, Pontus Skoglund, Kelly E. Graf, Mait Metspalu, Anders Albrechtsen, Ida Moltke, Simon Rasmussen, Thomas W. Stafford Jr, Ludovic Orlando, Ene Metspalu, Monika Karmin, Kristiina Tambets, Siiri Rootsi, Reedik Mägi, Paula F. Campos, Elena Balanovska, Oleg Balanovsky, Elza Khusnutdinova, Sergey Litvinov, Ludmila P. Osipova, Sardana A. Fedorova, Mikhail I. Voevoda, Michael DeGiorgio, Thomas Sicheritz-Pontén, Søren Brunak, Svetlana Demeshchenko, Toomas Kivisild, Richard Villems, Rasmus Nielsen, Mattias Jakobsson & Eske Willerslev. 2014. Upper palaeolithic Siberian genome reveals dual ancestry of Native Americans. *Nature* 505. 87–91.
- Raghavan, Maanasa et al. 2015. Genomic evidence for the pleistocene and recent population history of native Americans. *Science* 349. 1–10.
- Rajasingham, V. R. 2014. Interrogation in Muöt. *Mon-Khmer Studies* 43(1). 103–123.
- Ramsey, Robert S. 1987. *The languages of China*. Princeton: Princeton University Press.
- Rassadin, V. I. 1997. Tofalarskij jazyk. In Viktorija N. Jarceva, Nikolaj Nikolaevich Kazanskij & Andrej Aleksandrovich Kibrik (eds.), *Jazyki mira, vol. 2: Tjurkskiejazyki*, 372–383. Moskva: Nauka.

- Ratliff, Martha. 2010. *Hmong-Mien language history* (Pacific linguistics 613). Canberra: Pacific Linguistics.
- Rawski, Evelyn S. 2015. *Early modern China and Northeast Asia. Cross-border perspectives*. Cambridge: Cambridge University Press.
- Rédei, Károly. 1999. Zu den uralisch-jukagirischen Sprachkontakten. *Finnisch-Ugrische Forschungen* 55(1-3). 1–58.
- Refsing, Kirsten. 1986. *The Ainu language. The morphosyntax of the Shizunai dialect*. Århus: Aarhus University Press.
- Reich, David. 2018. *Who we are and how we got here. Ancient DNA and the new science of the human past*. New York: Pantheon Books.
- Reiko, Aso. 2015. Hateruma Yaeyama grammar. In Heinrich Patrick, Shinsho Miyara & Michinori Shimoji (eds.), *Handbook of the Ryukyuan languages. History, structure, and use* (Handbooks of Japanese Language and Linguistics 11), 423–447. Berlin: De Gruyter Mouton.
- Reio, Thomas G. 2011. Curiosity. In Sam Goldstein & Jack A. Naglieri (eds.), *Encyclopedia of child behavior and development*, 453–454. New York: Springer.
- Rialland, Annie. 2009. The African lax question prosody: Its realisation and geographical distribution. *Lingua* 119. 928–949.
- Rice, Karen. 1989. *A grammar of Slave* (Mouton Grammar Library 5). Berlin: Mouton de Gruyter.
- Rimsky-Korsakoff, Svetlana. 1967. Soviet Dungan: The Chinese language of Central Asia: Alphabet, phonology, morphology. *Monumenta Serica* 26. 352–421.
- Rimsky-Korsakoff, Svetlana. 1994. Zhongya zu ya donggan minzu (zhongguo huimin) de tiyu, jingkouling, shunkouliu ji huayu. *Monumenta Serica* 42. 483–520.
- Robbeets, Martine. 2015. *Diachrony of verb morphology. Japanese and the Transeurasian languages* (Trends in linguistics. Studies and Monographs 291). Berlin: De Gruyter Mouton.
- Røed, Knut H. 2008. Genetic analyses reveal independent domestication origins of Eurasian reindeer. *Proceedings of the Royal Society B: Biological Sciences* 275(1645). 1849–1855.
- Róna-Tas, András. 1998. The reconstruction of Proto-Turkic and the genetic question. In Lars Johanson & Éva Á. Csató (eds.), *The Turkic languages* (Routledge Language Family Series), 66–80. London: Routledge.
- Roos, Martina Erica. 2000. *The Western Yugur (Yellow Yugur) language: Grammar, texts, vocabulary*. Leiden: University of Leiden. (Doctoral dissertation.)
- Rosol, Sara Nelson, Christoph & Jürgen Renn. 2017. In the machine room of the Anthropocene. *The Anthropocene Review* 4(1). 2–8.
- Ross, Claudia & Jing-heng Sheng Ma. 2006. *Modern Mandarin Chinese grammar. A practical guide*. London: Routledge.
- Ross, John. 1877. *Corean primer*. Shanghai: American Preybyterian Mission Press.
- Ross, Lee. 2014 [1987]. The problem of construal in social inference and social psychology. In Neil E. Grunberg, R. E. Nisbett, Judith Rodin & Jerome E. Singer (eds.), *A distinctive*

## References

- approach to psychological research: The influence of Stanley Schachter*, 118–151. New York: Psychology Press.
- Rossano, Federico, Penelope Brown & Stephen C. Levinson. 2009. Gaze, questioning, and culture. In Jack Sidnell (ed.), *Conversation analysis: Comparative perspectives* (Studies in Interactional Sociolinguistics 27), 187–249. Cambridge: Cambridge University Press.
- Rozycki, William. 1993. *A problem in Proto-Tungus reconstruction: The phoneme \*k'* - 12. 203–212.
- Rybatzki, Volker. 2003a. Intra-Mongolic taxonomy. In Juha Janhunen (ed.), *The Mongolic languages* (Routledge Language Family Series), 364–390. London: Routledge.
- Rybatzki, Volker. 2003b. Middle Mongol. In Juha Janhunen (ed.), *The Mongolic languages* (Routledge Language Family Series), 57–82. London: Routledge.
- Sadock, Jerrold M. 1984. West Greenlandic. In Williams S. Chisolm (ed.), *Interrogativity. A colloquium on the grammar, typology and pragmatics of questions in seven diverse languages, Cleveland, Ohio, October 5th 1981-May 3rd 1982* (Typological Studies in Language 4), 189–214. Amsterdam: Benjamins.
- Sadock, Jerrold M. & Arnold M. Zwicky. 1985. Speech act distinctions in syntax. In Timothy Shopen (ed.), *Language typology and syntactic description, vol 1: Clause structure*, 155–196. Cambridge: Cambridge University Press.
- Safonova, Tatiana & István Sántha. 2013. *Culture contact in Evenki land. A cybernetic anthropology of the Baikal region* (Inner Asia Book Series 7). Leiden: Global Oriental.
- Sagart, Laurent. 2016. The wider connections of Austronesian. A response to Blust 2009. *Diachronica* 33(2). 255–281.
- Saltzman, Moira. 2014. *Language contact and morphological change in Jéjueo*. Detroit: Wayne State University. (Master's thesis.)
- Sammallahti, Pekka. 1988. Historical phonology of the Uralic languages. In Denis Sinor (ed.), *The Uralic languages* (Handbuch der Orientalistik 8(1)), 478–554. Leiden: Brill.
- San Roque, Lila, Simeon Floyd & Elisabeth Norcliffe. 2017. Evidentiality and interrogativity. *Lingua* 186/187. 120–143.
- Sanada, Shinji & Yukio Uemura. 2007. Japanese dialects and Ryukyuan. In Miyaoka Osahito, Sakiyama Osamu & Michael E. Krauss (eds.), *The vanishing languages of the Pacific Rim* (Oxford Linguistics), 355–367. Oxford: Oxford University Press.
- Sandman, Erika. 2012. Bonan grammatical features in Wutun Mandarin. *Mémoires de la Société Finno-Ougrienne* 264. 375–387.
- Sandman, Erika. 2016. *A grammar of Wutun*. Helsinki: University of Helsinki. (Doctoral dissertation.)
- Sandman, Erika & Camille Simon. 2016. Tibetan as a “model language” in the Amdo Sprachbund: Evidence from Salar and Wutun. *Journal of South Asian Languages and Linguistics* 3(1). 85–122.
- Sanitt, Nigel. 2007. A mingled yarn: Problematology and science. *Revue Internationale de Philosophie* 61(242-4). 435–449.
- Sanitt, Nigel. 2011. Science and language. *Language Sciences* 33. 559–561.
- Sankararaman, Sriram, Nick Patterson & David Reich. 2016. The combined landscape of Denisovan and Neanderthal ancestry in present-day humans. *Current Biology* 26. 1–7.

- Schiefner, Anton. 1871. Über Baron Gerhard von Maydell's jukagirische Sprachproben. *Bulletin de l'Académie Impériale des Sciences de St Pétersbourg* 15. 86–103.
- Schiefner, Anton. 1874. Baron Gerhard von Maydell's tungusische Sprachproben. *Bulletin de l'Académie Impériale des Sciences de St Pétersbourg* 20(2). 210–246.
- Schmalz, Mark. 2012. Towards a full description of the focus system in Tundra Yukaghir. *Linguistic Discovery* 10(2). 53–108.
- Schmalz, Mark. 2013. *Aspects of the grammar of Tundra Yukaghir*. Amsterdam: University of Amsterdam. (Doctoral dissertation.)
- Schmidt, Peter. 1923a. The language of the Negidals. *Acta Universitatis Latviensis* 5. 3–38.
- Schmidt, Peter. 1923b. The language of the Olchas. *Acta Universitatis Latviensis* 8. 229–288.
- Schmidt, Peter. 1928a. The language of the Oroches. *Acta Universitatis Latviensis* 17. 17–62.
- Schmidt, Peter. 1928b. The language of the Samagirs. *Acta Universitatis Latviensis* 19. 219–49.
- Schmitt, Rüdiger. 2000. *Die iranischen Sprachen in Geschichte und Gegenwart*. Wiesbaden: Reichert Verlag.
- Schönig, Claus. 1993. Das Tofa, eine neue türkische Schriftsprache in der Sowjetunion. *Ural-Altäische Jahrbücher* 12. 192–201.
- Schönig, Claus. 1998. South Siberian Turkic. In Lars Johanson & Éva Á. Csató (eds.), *The Turkic languages* (Routledge Language Family Series), 403–416. London: Routledge.
- Schönig, Claus. 1999. The internal division of modern Turkic languages and its historical implications. *Acta Orientalia Academiae Scientiarum Hungaricae* 52(1). 63–95.
- Schönig, Claus. 2003. Turko-Mongolic relations. In Juha Janhunen (ed.), *The Mongolic languages* (Routledge Language Family Series), 403–419. London: Routledge.
- Schröder, Dominik. 1964. Der Dialekt der Monguor. In B. Spuler (ed.), *Mongolistik* (Handbuch der Orientalistik 1(5.2)), 143–158. Leiden: Brill.
- Schulze, Wolfgang. 2007. Communication or memory mismatch: Towards a cognitive typology of questions. In Günter Radden, Klaus-Michael Köpcke, Thomas Berg & Peter Siemund (eds.), *Aspects of meaning construction*, 247–264. Amsterdam: Benjamins.
- Sean, Lee. 2015. A sketch of language history in the Korean Peninsula. *PlosOne* 10(5). 1–12.
- Sean, Lee & Toshikazu Hasegawa. 2013. Evolution of the Ainu language in space and time. *PlosOne* 10(5). 1–6.
- Sean, Lee & Hasegawa Toshikazu. 2011. Bayesian phylogenetic analysis supports an agricultural origin of Japonic languages. *Proceedings of the Royal Society B: Biological Sciences* 278. 3662–3669.
- Sechenbaatar, Borjigin. 2003. *The Chakhar dialect of Mongol. A morphological description* (Mémoire de la Société Finno-Ougrienne 243). Helsinki: The Finno-Ugrian Society.
- Seebold, Elmar. 2002. *Kluge. Etymologisches Wörterbuch der deutschen Sprache*. 24th edn. Berlin: de Gruyter.
- Sekerina, Irina A. 1994. Copper Island (Mednyj) Aleut CIA: A mixed language. *Languages of the World* 8. 14–31.

## References

- Sem, Lidija I. 1976. *Očerki dialektov nanajnskogo jazyka. bikinskij (ussuriskij) dialekt*. Leningrad: Nauka.
- Shapiro, Roman. 2010. Chinese Pidgin Russian. *Journal of Pidgin and Creole Languages* 25(1). 5–62.
- Shaw, Robert B. S. 1876. *On the Ghalchah languages (Wakhí and Sariḳolí)*. Calcutta: C. B. Lewis, Baptist Mission Press.
- Shaw, Robert B. S. 1878. *A sketch of the Turki language of Eastern Turkestan (Kàshgar and Yarkand)*. Calcutta: Baptist Mission Press.
- Shevelov, George Y. 1993. Ukrainian. In Bernard Comrie & Greville G. Corbett (eds.), *The Slavonic languages* (Routledge Language Family Series), 947–998. London: Routledge.
- Shi Feng. 2009. Shiquan fangyan de yiwen daici. *Ankang xueyuan xuebao* 21(2). 14–17.
- Shibatani, Masayoshi. 1990. *The languages of Japan* (Cambridge Language Surveys). Cambridge: Cambridge University Press.
- Shigeno, Hirumi. 2010. Ura (Amami Ryukyuan). In Thomas Pellard & Michinori Shimoji (eds.), *An introduction to Ryukyuan languages*, 15–34. Tōkyō: Research Institute for Languages, Cultures of Asia & Africa.
- Shimoji, Michinori. 2010. Ryukyuan languages: An introduction. In Thomas Pellard & Michinori Shimoji (eds.), *An introduction to Ryukyuan languages*, 1–13. Tōkyō: Research Institute for Languages, Cultures of Asia & Africa.
- Shimoji, Michinori. 2011a. Irabu Ryukyuan. In Yasuhiro Yamakoshi (ed.), *Grammatical sketches from the field*, 79–136. Tōkyō: Research Institute for Languages, Cultures of Asia & Africa.
- Shimoji, Michinori. 2011b. Quasi-kakari musubi in Irabu. *Japanese/Korean linguistics* 18. 114–125.
- Shimuněk, Andrew. 2014. A new decipherment and linguistic reconstruction of the Kitan –Chinese bilingual inscription of 1134 A.D. *Acta Orientalia Academiae Scientiarum Hungaricae* 67(1). 97–118.
- Shimuněk, Andrew. 2017. *Languages of ancient southern Mongolia and north China. A historical-comparative study of the Serbi or Xianbei branch of the Serbi-Mongolic language family, with an analysis of northeastern frontier Chinese and Old Tibetan phonology* (Tunguso Sibirica 40). Wiesbaden: Harrassowitz.
- Shimuněk, Andrew, Christopher I. Beckwith, Jonathan North Washington, Nicholas Kontovas & Kurban Niyaz. 2015. The earliest attested Turkic language. The Chieh 𐰉 (\*Kir) language of the fourth century A.D. *Journal Asiatique* 303(1). 143–151.
- Shinzato, Rumiko. 2015. Okinawan kakari musubi in historical and comparative perspectives. In Heinrich Patrick, Shinsho Miyara & Michinori Shimoji (eds.), *Handbook of the Ryukyuan languages. History, structure, and use* (Handbooks of Japanese Language and Linguistics 11), 299–320. Berlin: De Gruyter Mouton.
- Shinzato, Rumiko & Leon A. Serafim. 2013. *Synchrony and diachrony of Okinawan kakari musubi in comparative perspective with premodern Japanese* (Languages of Asia). Leiden: Brill.
- Shiraishi, Hidetoshi. 2006. *Topics in Nivkh phonology*. Groningen: Rijksuniversiteit Groningen. (Doctoral dissertation.)

- Shiraishi, Hidetoshi & Itsuji Tangiku. 2013. Nivufugo amūru hōgen no kiso goi. *Hoppō-gengokenkyū* 3. 201–212.
- Sicoli, Mark A. & Gary Holton. 2014. Linguistic phylogenies support back-migration from Beringia to Asia. *PlosOne* 9(3). 1–8.
- Sicoli, Mark A., Tanya Stivers, Nicholas J. Enfield & Stephen C. Levinson. 2014. Marked initial pitch in questions signals marked communicative function. *Language and Speech* 58(2). 1–20.
- Sieg, Emil & Wilhelm Siegling. 1931. *Tocharische Grammatik* (Göttinger Sammlung indogermanischer Grammatiken). Göttingen: Vandenhoeck & Ruprecht.
- Siegl, Florian. 2012. Yes/no questions and the interrogative mood in Forest Enets. *Mémoires de la Société Finno-Ougrienne* 264. 399–408.
- Siegl, Florian. 2013. *Materials of Forest Enets. An indigenous language of northern Siberia* (Mémoires de la Société Finno-Ougrienne 267). Tartu: Tartu University Press.
- Siemund, Peter. 2001. Interrogative constructions. In Martin Haspelmath, Ekkehard König, Wulf Oesterreicher & Wolfgang Raible (eds.), *Language typology and language universals. An international handbook* (Handbücher zur Sprach- und Kommunikationswissenschaft 20(2)), 1010–1028. Berlin: De Gruyter Mouton.
- Simon, Camille. 2015. Linguistic evidence of Salar-Tibetan contacts in Amdo. In Marie-Paule Hille, Bianca Horlemann & Paul K. Nietupski (eds.), *Muslims in Amdo Tibetan society. Multidisciplinary perspectives* (Studies in Modern Tibetan Culture), 87–107. Lanham: Lexington Books.
- Sinha, Chris. 2013. Niche construction, too, unifies praxis and symbolization. *Language and Cognition* 5(2/3). 261–271.
- Sinor, Denis (ed.). 1988. *The Uralic languages* (Handbuch der Orientalistik 8(1)). Leiden: Brill.
- Sinor, Denis. 1990. Introduction: The concept of Inner Asia. In Denis Sinor (ed.), *The Cambridge History of early Inner Asia*, 1–18. Cambridge: Cambridge University Press.
- Siqinchaoketu. 1999. *Kangjiayu*. Shanghai: Shanghai yuandong chubanshe.
- Siqinchaoketu. 2002. Kangjiayu gaikuang. *Minzu yuwen* 6. 66–77.
- Siska, Veronika, Eppie R. Jones, Jeon Sungwon, Bhak Youngjune, Kim Hak-Min, Cho Yun Sung, Kim Hyunho, Lee Kyusang, Elizaveta Veselovskaya, Tatiana Balueva, Marcos Gallego-Llorente, Michael Hofreiter, Daniel G. Bradley, Anders Eriksson, Ron Pinhasi, Bhak Jong & Andrea Manica. 2017. Genome-wide data from two early Neolithic East Asian individuals dating to 7700 years ago. *Science Advances* 3. 1–10.
- Skoglund, Pontus, Swapan Mallick, Maria Cátira Bortolini, Niru Chennagiri, Tábita Hünemeier, Maria Luíza Petzl-Erler, Francisco Mauro Salzano, Nick Patterson & David Reich. 2015. Genetic evidence for two founding populations of the Americas. *Nature* 525. 104–108.
- Skribnik, Elena. 2003. Buryat. In Juha Janhunen (ed.), *The Mongolic languages* (Routledge Language Family Series), 102–128. London: Routledge.
- Skribnik, Elena. 2004. Sprachen Sibiriens. Sprachkontakte und Sprachgeschichte, Sprachgeburts und Sprachtod. In Peter Schrijver & Peter-Arnold Mumm (eds.), *Sprachgeburts und Sprachtod*, 141–170. Bremen: Hempen Verlag.

## References

- Slater, Keith W. 2003a. *A grammar of Mangghuer* (Routledge Studies in Asian Linguistics). London: Routledge Curzon.
- Slater, Keith W. 2003b. Mangghuer. In Juha Janhunen (ed.), *The Mongolic languages* (Routledge Language Family Series), 307–324. London: Routledge.
- Sohn, Ho-Min. 1994. *Korean* (Descriptive grammars). London: Routledge.
- Sohn, Ho-Min. 1999. *The Korean language* (Cambridge Language Surveys). Cambridge: Cambridge University Press.
- Sohn, Ho-Min. 2012. Middle Korean. In Nicolas Tranter (ed.), *The languages of Japan and Korea* (Routledge Language Family Series), 73–122. London: Routledge.
- Sohn, Ho-Min. 2015. Middle Korean and Pre-Modern Korean. In Lucien Brown & Yeon Jaehoon (eds.), *The handbook of Korean linguistics*, 439–458. Malden: Wiley Blackwell.
- Song, Jae Jung. 2005. *The Korean language. Structure, use and context*. London: Routledge.
- Sotavalta, Arvo. 1978. Westlamutische Materialien. *Mémoires de la Société Finno-Ougrienne* 168. 1–212.
- Spencer, Andrew. 1995. Incorporation in Chukchi. *Language* 71(3). 439–489.
- Stachowski, Marek. 1990. Der Lautwandel in der Deklination des Jakutischen Interrogativpronomens tuox ‘was?’ *Zeszyty Naukowe UJ. Prace Językoznawcze* 101. 115–120.
- Stachowski, Marek. 1993. *Dolganischer Wortschatz*. Krakow: Nakladem Univ. Jagiellońskiego.
- Stachowski, Marek. 2015. Turkic pronouns against a Uralic background. *Iran and the Caucasus* 19. 79–86.
- Stachowski, Marek & Astrid Menz. 1998. Yakut. In Lars Johanson & Éva Á. Csató (eds.), *The Turkic languages* (Routledge Language Family Series), 417–433. London: Routledge.
- Steffensen, Sune Vork & Alwin Fill. 2014. Ecolinguistics: The state of the art and future horizons. *Language Sciences* 41. 6–25.
- Stern, Dieter. 2005. Taimyr Pidgin Russian (Govorka). *Russian Linguistics* 29(3). 289–318.
- Stern, Dieter. 2012. *Tajmyr-Pidgin-Russisch. Kolonialer Sprachkontakt in Nordsibirien* (Studies on language and culture in Central and Eastern Europe 19). Munich: Otto Sagner.
- Stilo, Donald L. 2015. An introduction to the Atlas of the Araxes-Iran linguistic area. In Bernard Comrie & Lucía Golluscio (eds.), *Language contact and documentation*, 343–355. Berlin: Mouton de Gruyter.
- Stivers, Tanya, Nicholas J. Enfield, Penelope Brown, Christina Englert, Hayashi Makoto, Trine Heinemann, Gertie Hoymann, Federico Rossano, Jan Peter de Ruiter, Yoon Kyung-Eun & Stephen C. Levinson. 2009. Universals and cultural variation in turn-taking in conversation. *Proceedings of the National Academy of Sciences* 106(26). 10587–10592.
- Stolz, Thomas, Nataliya Levkovych, Aina Urdze, Julia Nintemann & Maja Robbers. 2017. *Spatial interrogatives in Europe and beyond. Where, whither, whence* (Studia Typologica 20). Berlin: De Gruyter Mouton.
- Street, John C. 2008/09. The interrogative particle in early Middle Mongolian. *Mongolian Studies* 30/31. 43–82.

- Sulamo, Dagnachew Degu. 2013. *Interrogative constructions in Hadiyya*. Saarbrücken: Lambert Academic Publishing.
- Sumner, Francis B. 1922. The organism and its environment. *The Scientific Monthly* 14(3). 223–233.
- Sun Chaofen. 2006. *Chinese. A linguistic introduction*. Cambridge: Cambridge University Press.
- Sun Hongkai, Liu Guangkun, Li Fengxiang, Ela Thurgood & Graham Thurgood. 2009. *A grammar of Anong. Language death under intense contact* (Brill's Tibetan Studies Library. Languages of the Greater Himalayan Region 5(9)). Leiden: Brill.
- Sun Hongkai, [Katia Chirkova] Qi Kajia & Liu Guangkun. 1996. *Baimayu yanjiu*. Peking: Minzu chubanshe.
- Sun Qinglin. 2015. Shanxi yanggao fangyan de yiwen daici. *Mingzuo xinshang* 21. 150–152.
- Sun, Jackson T.-S. 1986. *Aspects of the phonology of Amdo Tibetan: Ndzorge s̄æme xvra dialect* (Monumenta Serindica 16). Tōkyō: Institute for the Study of Languages, Cultures of Asia & Africa.
- Sun, Jackson T.-S. 1993. Evidentials in Amdo Tibetan. *Bulletin of the Institute of History and Philology* 63(4). 945–1001.
- Sun, Jackson T.-S. 2003a. Caodeng rGyalrong. In Graham Thurgood & Randy L. LaPolla (eds.), *The Sino-Tibetan languages* (Routledge Language Family Series), 490–502. London: Routledge.
- Sun, Jackson T.-S. 2003b. Phonological profile of Zhongu: A new Tibetan dialect of northern Sichuan. *Language and Linguistics* 4(4). 769–836.
- Sun, Jackson T.-S. 2006. Special linguistic features of gSerpa Tibetan. *Linguistics of the Tibeto-Burman Area* 29(1). 107–125.
- Sun, Jackson T.-S. 2007. Perfective stem renovation in Khalong Tibetan. In Roland Bielmeier & Felix Haller (eds.), *Linguistics of the Himalayas and beyond*, 323–340. Berlin: Mouton de Gruyter.
- Sunik, Orest. 1958. *Kur-urmijskij dialekt*. Leningrad: Nauka.
- Sussex, Roland & Paul Cubberley. 2006. *The Slavic languages* (Cambridge Language Surveys). Cambridge: Cambridge University Press.
- Suzuki, Hiroyuki. 2006. Amudo chibettogo chū aba [rNgawa] hōgen no onsei bunseki. *Asian and African Languages and Linguistics* 1. 59–88.
- Svantesson, Jan-Olof. 2003. Khalkha. In Juha Janhunen (ed.), *The Mongolic languages* (Routledge Language Family Series), 154–176. London: Routledge.
- Svantesson, Jan-Olof, Anna Tsedina, Anastasia Karlsson & Vivian Franzén. 2005. *The phonology of Mongolian* (Oxford Linguistics. The Phonology of the World's Languages). Oxford: Oxford University Press.
- Swenson, Rod & Michael T. Turvey. 1991. Thermodynamic reasons for perception-action cycles. *Ecological Psychology* 3(4). 317–348.
- Szokolszky, Agnes. 2013. Interview with Ulric Neisser. *Ecological Psychology* 25. 182–199.
- Taaffe, Robert N. 1990. The geographical setting. In Denis Sinor (ed.), *The Cambridge history of early Inner Asia*, 19–40. Cambridge: Cambridge University Press.

## References

- Takahashi, Yasushige. 2013. Ainugo tokachi hōgen ni okeru shōkosei to jojutsu ruikei. *Hoppō gengo kenkyū* 3. 129–136.
- Takehiro, Sato, Amano Tetsuya, Ono Hiroko, Ishida Hajime, Kodera Haruto, Matsumura Hirofumi, Yoneda Minoru & Masuda Ryuichi. 2007. Origins and genetic features of the Okhotsk people, revealed by ancient mitochondrial DNA analysis. *Journal of Human Genetics* 52. 618–627.
- Tamura, Suzuko. 2000. *The Ainu language* (ICHEL Linguistic Studies 2). Tōkyō: Sanseido.
- Taylor, Archer. 1943. The riddle. *California Folklore Quarterly* 2(2). 129–147.
- Thomason, Sarah G. 2001. *Language contact. An introduction*. Edinburgh: Edinburgh University Press.
- Thompson, Hanne-Ruth. 2012. *Bengali* (London Oriental and African Language Library 18). Amsterdam: Benjamins.
- Tittel, Hans (transl.) 1922. Die Sprache der Ainus von Sachalin. *Mitteilungen der Deutschen Gesellschaft für Natur- und Völkerkunde Ostasiens* 17. 70–89.
- Todaeva, Buljas Ch. 1959. Über die Sprache der Tung-hsiang. *Acta Orientalia Academiae Scientiarum Hungaricae* 9(3). 273–310.
- Todaeva, Buljas Ch. 1963. Einige Besonderheiten der Paoan-Sprache. *Acta Orientalia Academiae Scientiarum Hungaricae* 16(2). 175–197.
- Tolskaya, Maria & Inna Tolskaya. 2008. Question particles or what? Open alternative questions in Manchu-Tungusic languages. In Edward J. Vajda (ed.), *Subordination and coordination strategies in North Asian languages* (Current Issues in Linguistic Theory Series 300), 97–109. Amsterdam: Benjamins.
- Tomasello, Michael. 2003. *Constructing a language. A usage-based theory of language acquisition*. Cambridge: Harvard University Press.
- Tomasello, Michael. 2008. *Origins of human communication*. Cambridge: The MIT Press.
- Tomasello, Michael. 2014a. *A natural history of human thinking*. Cambridge: Harvard University Press.
- Tomasello, Michael. 2014b. The ultra-social animal. *European Journal of Social Psychology* 44. 187–194.
- Tomomi, Sato. 2002. A basic vocabulary of the Samani dialect of Ainu. *Hokkaidō daigaku bungaku kenkyū-ka kiyō* 106. 91–126.
- Tooru, Hayasi, Sabit Rozi, Tahirjan Muhämmät & Wang Jianxin. 1999. *A Šäyxil vocabulary. A preliminary report of linguistic research in Šäyxil village, southwestern Xinjiang* (Contribution to the Studies of Eurasian Languages Series 2). Kyōto: Kyōto University.
- Tournadre, Nicolas. 2005. L'aire linguistique tibétaine et ses divers dialectes. *Lalies* 25. 7–56.
- Tournadre, Nicolas. 2014. The Tibetic languages and their classification. In Nathan W. Hill & Thomas Owen-Smith (eds.), *Trans-Himalayan linguistics. Historical and descriptive linguistics of the Himalayan area* (Trends in Linguistics. Studies and Monographs 266), 105–129. Berlin: De Gruyter Mouton.
- Tournadre, Nicolas & Randy LaPolla. 2014. Towards a new approach to evidentiality. Issues and directions for research. *Linguistics of the Tibeto-Burman area* 37(2). 240–263.

- Tranter, Nicolas. 2012a. Classical Japanese. In Nicolas Tranter (ed.), *The languages of Japan and Korea* (Routledge Language Family Series), 212–245. London: Routledge.
- Tranter, Nicolas. 2012b. Introduction: Typology and area in Japan and Korea. In Nicolas Tranter (ed.), *The languages of Japan and Korea* (Routledge Language Family Series), 3–23. London: Routledge.
- Tranter, Nicolas & Mika Kizu. 2012. Modern Japanese. In Nicolas Tranter (ed.), *The languages of Japan and Korea* (Routledge Language Family Series), 268–312. London: Routledge.
- Treis, Yvonne. 2014. Interrogativity in Baskeet. *Abhandlungen für die Kunde des Morgenlandes* 91. 41–78.
- Trudgill, Peter. 2011. *Sociolinguistic typology. Social determinants of linguistic complexity*. Oxford: Oxford University Press.
- Tsumagari, Toshiro. 1997. Linguistic diversity and national borders of Tungusic. *Senri Ethnological Studies* 44. 175–186.
- Tsumagari, Toshiro. 2003. Dagur. In Juha Janhunen (ed.), *The Mongolic languages* (Routledge Language Family Series), 129–153. London: Routledge.
- Tsumagari, Toshiro. 2009a. A sketch of Solon grammar. *Journal of the Center for Northern Humanities* 2. 1–21.
- Tsumagari, Toshiro. 2009b. Grammatical outline of Uilta (revised). *Journal of the Graduate School of Letters* 4. 1–21.
- Tsumagari, Toshiro, Megumi Kurebito & Fubito Endo. 2007. Siberia: Tungusic and Palaeosiberian. In Miyaoka Osahito, Sakiyama Osamu & Michael E. Krauss (eds.), *The vanishing languages of the Pacific Rim* (Oxford Linguistics), 387–405. Oxford: Oxford University Press.
- Tuohuti Litifu. 2012. *Xiandai weiwu'eryu cankao yufa*. Peking: Zhongguo shehui kexue chubanshe.
- Turvey, Michael T. 2009. On the notion and implications of organism-environment system. *Ecological Psychology* 21(2). 97–111.
- Uemura, Yukio. 2003. *The Ryukyuan language*. Kyōto: Nakanishi.
- Ultan, Russell. 1978. Some general characteristics of interrogative systems. In Joseph Greenberg (ed.), *Universals of human language, vol. 4: Syntax*, 211–248. Stanford: Stanford University Press.
- Vajda, Edward J. 2004. *Ket* (Languages of the World/Materials 204). Munich: Lincom Europa.
- Vajda, Edward J. 2009a. Loanwords in Ket. In Martin Haspelmath & Uri Tadmor (eds.), *Loanwords in the world's languages. A comparative handbook*, 471–495. Berlin: De Gruyter.
- Vajda, Edward J. 2009b. The languages of Siberia. *Language and Linguistics Compass* 3(1). 424–440.
- Vajda, Edward J. 2010. A Siberian link with Na-Dene languages. *Anthropological Papers of the University of Alaska* 5. 33–99.
- Vajda, Edward J. 2013. Vestigial possessive morphology in Na-Dene and Yeniseian. *Alaska Native Language Center Working Papers* 11. 79–91.

## References

- Vakhtin, Nikolai. 1998. Copper Island Aleut: A case of language “resurrection”. In Lenore A. Grenoble & Lindsay J. Whaley (eds.), *Endangered languages. Current issues and further prospects*, 317–327. Cambridge: Cambridge University Press.
- Vakhtin, Nikolai. 2000. *Jazyk sireniskix èskimosov: Teksty, grammaticheskie i slovarnye materialy* (Lincom Studies in Asian Linguistics 33). Munich: Lincom Europa.
- Van Alsenoy, Lauren & Johan van der Auwera. 2015. Indefinite pronouns in Uralic languages. In Matti Miestamo, Anne Tamm & Beáta Wagner-Nagy (eds.), *Negation in Uralic Languages* (Typological Studies in Language 108), 519–546. Amsterdam: Benjamins.
- van den Berg, René & Robert L. Busenitz. 2012. *A grammar of Balantak. A language of Eastern Sulawesi* (SIL eBook 40). Dallas: SIL International.
- van der Auwera, Johan & Dmitry Idiatov. 2008. From nominalizations to questions. Evidence from Tucanoan. *Linguistics (Jezikoslovlje)* 9(1-2). 35–51.
- van der Auwera, Johan & Jan Nuyts. 2007. Cognitive linguistics and linguistic typology. In Dirk Geeraerts & Hubert Cuyckens (eds.), *The Oxford handbook of Cognitive Linguistics*, 1074–1091. Oxford: Oxford University Press.
- van der Lubbe, Gijs & Akiko Tokunaga. 2015. Okinoerabu grammar. In Heinrich Patrick, Shinsho Miyara & Michinori Shimoji (eds.), *Handbook of the Ryukyuan languages. History, structure, and use* (Handbooks of Japanese Language and Linguistics 11), 345–377. Berlin: De Gruyter Mouton.
- van Driem, George. 2014. Trans-Himalayan. In Nathan Hill & Thomas Owen-Smith (eds.), *Trans-Himalayan linguistics. Historical and descriptive linguistics of the Himalayan area* (Trends in Linguistics. Studies and Monographs 266), 11–40. Berlin: Mouton de Gruyter.
- Veluppilai, Viveka. 2012. *An introduction to linguistic typology*. Amsterdam: Benjamins.
- Voegelin, Carl F. & Florence M. Voegelin. 1964. Languages of the world. Native America fascicle one. *Anthropological Linguistics* 6(6). 1–149.
- Volodin, Aleksandr P. 2001. Imperative in Kerek. In Viktor S. Khrakovski (ed.), *Typology of imperative constructions* (Lincom Studies in Theoretical Linguistics 9), 145–158. Munich: Lincom Europa.
- Volodko, Natalia V., Elena B. Starikovskaya, Ilya O. Mazunin, Nikolai P. Eltsov, Polina V. Naidenko, Douglas C. Wallace & Rem I. Sukernik. 2008. Mitochondrial genome diversity in Arctic Siberians, with particular reference to the evolutionary history of Beringia and Pleistocenic peopling of the Americas. *The American Journal of Human Genetics* 82. 1084–1100.
- von Stumm, Sophie, Benedikt Hell & Tomas Chamorro-Premuzic. 2011. The hungry mind: Intellectual curiosity is the third pillar of academic performance. *Perspectives on Psychological Science* 48(6(6)). 574–588.
- von Möllendorff, Paul G. 1892. *A Manchu grammar with analysed texts*. Shanghai: American Presbyterian Mission Press.
- Vovin, Alexander. 1993. *A reconstruction of Proto-Ainu* (Brill’s Japanese Studies Library 4). Leiden: Brill.

- Vovin, Alexander. 2004. Some thoughts on the origins of the Old Turkic 12-year animal cycle. *Central Asiatic Journal* 48(1). 118–132.
- Vovin, Alexander. 2005. *A descriptive comparative grammar of Western Old Japanese, vol. 1: Phonology, script, lexicon and nominals* (Languages of Asia 3). Folkestone: Global Oriental.
- Vovin, Alexander. 2009. *A descriptive comparative grammar of Western Old Japanese, vol. 2: Adjectives, verbs, conjunctions, particles, postpositions, indexes* (Languages of Asia 8). Folkestone: Global Oriental.
- Vovin, Alexander. 2013a. From Koguryō to T’amna. Slowly riding to the South with speakers of Proto-Korean. *Korean Linguistics* 15(2). 222–240.
- Vovin, Alexander. 2013b. Northeastern and Central Asia: “Altaic” linguistic history. In Peter Bellwood (ed.), *The global prehistory of human migration*, 197–208. Malden: Wiley Blackwell.
- Vovin, Alexander. 2015. Eskimo loanwords in northern Tungusic. *Iran and the Caucasus* 19. 87–95.
- Vovin, Alexander. 2016. On the linguistic prehistory of Hokkaidō. *Studia Orientalia* 117. 29–38.
- Vovin, Alexander. 2017. *Interpretation of the Hüis Tolgoi inscription*. Paper presented at PIAC, 31 August 2017, Székesfehérvár.
- Vovin, Alexander, Edward Vajda & Étienne de la Vaissière. 2016. Who were the \*kjet (羯) and what language did they speak? *Journal Asiatique* 304(1). 125–144.
- Wade, Terence. 2011. *A comprehensive Russian grammar* (Blackwell Reference Grammars 8). Malden: Wiley-Blackwell.
- Wagner-Nagy, Beáta. 2015. Negation in Selkup. In Matti Miestamo, Anne Tamm & Beáta Wagner-Nagy (eds.), *Negation in Uralic languages* (Typological Studies in Language 108), 133–158. Amsterdam: Benjamins.
- Wagner-Nagy, Beáta. 2016. Nenets. In Peter O. Müller, Ingeborg Ohnheiser, Susan Olsen & Franz Rainer (eds.), *Word-formation: An international handbook of the languages of Europe* (Handbücher zur Sprach- und Kommunikationswissenschaft 40(5)), 3197–3209. Berlin: De Gruyter Mouton.
- Wang Qingfeng. 2005. *Manyu yanjiu*. Peking: Minzu chubanshe.
- Watters, David E. 2002. *A grammar of Kham* (Cambridge Grammatical Descriptions). Cambridge: Cambridge University Press.
- Watters, David E. 2006. Notes on Kusunda grammar. A language isolate of Nepal. *Himalayan Linguistic Archive* 3. 1–182.
- Weiers, Michael. 1972. *Die Sprache der Moghol der Provinz Herat in Afghanistan: Sprachmaterial, Grammatik, Wortliste*. Opladen: Westdeutscher Verlag.
- Weiqiang, Hei & Leng Yongliang. 2013. Shanbei suide fangyan de yiwen daici. *Ningbo daxue xueyuan (renwen kexue ban)* 26(2). 51–54.
- Welsch, Wolfgang. 2012. *Mensch und Welt. Eine evolutionäre Perspektive der Philosophie*. Munich: Beck.
- Werner, Heinrich. 1995. *Zur Typologie der Jenissej-Sprachen* (Veröffentlichungen der Societas Uralo-Altaica 45). Wiesbaden: Harrassowitz.

## References

- Werner, Heinrich. 1997a. *Abriß der kottischen Grammatik* (Tunguso Sibirica 4). Wiesbaden: Harrassowitz.
- Werner, Heinrich. 1997b. *Das Jugische (Sym-Ketische)* (Veröffentlichungen der Societas Uralo-Altaica 50). Wiesbaden: Harrassowitz.
- Werner, Heinrich. 1997c. *Die ketische Sprache* (Tunguso Sibirica 3). Wiesbaden: Harrassowitz.
- Werner, Heinrich. 2004. *Zur jenseitsjisch-indianischen Urverwandtschaft*. Wiesbaden: Harrassowitz.
- Whaley, Lindsay J. 2005. *An inflecting question particle in Oroqen*. Paper presented at Association for Linguistic Typology, 21–25 July 2005, Padang.
- Whaley, Lindsay J. & Fengxiang Li. 2000. Oroqen dialects. *Central Asiatic Journal* 44. 105–130.
- Whitman, John. 2012. The relationship between Japanese and Korean. In Nicolas Tranter (ed.), *The languages of Japan and Korea* (Routledge Language Family Series), 24–38. London: Routledge.
- Whitman, John. 2015. Old Korean. In Lucien Brown & Yeon Jaehoon (eds.), *The handbook of Korean linguistics*, 421–438. Malden: Wiley Blackwell.
- Wiener, Gerald. 2013. The yak, an essential element of the high altitude regions of Central Asia. *Études mongoles et sibériennes, centrasiatiques et tibétaines* 43/44. 1–10.
- Wilbur, Joshua. 2014. *A grammar of Pite Saami* (Studies in Diversity Linguistics 5). Berlin: Language Science Press.
- Windfuhr, Gernot & John R. Perry. 2009. Persian and Tajik. In Gernot Windfuhr (ed.), *The Iranian languages* (Routledge Language Family Series), 416–544. London: Routledge.
- Winter, Werner. 1998. Tocharian. In Giacalone Anna Ramat & Paolo Ramat (eds.), *The Indo-European languages* (Routledge Language Family Series), 154–168. London: Routledge.
- Witsen, Nicolaas. 1705. *Noord en Oost Tartarye*. Amsterdam: François Halma.
- Wu Hongwei. 1999. *Tuwayu*. Shanghai: Shanghai yuandong chubanshe.
- Wuge Shouping & Cheng Mingyuan. 1730. *Manju nikan hergen-i cing wen ki meng bithe. Manhan zi qingwen qimeng*. 4 vols.
- Wurm, Stephen A. 1997. Two Turkic-based hybrid languages in northwestern China. *Turkic Languages* 1. 241–253.
- Wylie, Alexander. 1855. *Translation of the Ts'ing wan k'e mung, a Chinese grammar of the Manchu Tartar language; with introductory notes on Manchu literature*. Shanghai: London Mission Press.
- Xie Xiaolan & Zhang Shumin. 1990. Gansu linxia fangyan de yiwenju. *Lanzhou daxuexuebao* 18(3). 141–146.
- Xiren Kuerban & Alimujiang Xiren. 2015. *Tajikeyu 366 ju huihua ju*. Peking: Shehui kexue wenxian chubanshe.
- Xu, Dan. 2006. *Typological change in Chinese syntax* (Oxford Linguistics). Oxford: Oxford University Press.
- Xu, Dan. 2014. *Tangwanghua yanjiu*. Peking: Minzu chubanshe.

- Xuan Dewu, Jin Xiangyuan & Zhao Xi. 1985. *Chaoxianyu jianzhi*. Peking: Minzu chubanshe.
- Yamada, Masahiro, Thomas Pellard & Michinori Shimoji. 2015. Dunan grammar (Yonaguni Ryukyuan). In Heinrich Patrick, Shinsho Miyara & Michinori Shimoji (eds.), *Handbook of the Ryukyuan languages. History, structure, and use* (Handbooks of Japanese Language and Linguistics 11), 449–478. Berlin: De Gruyter Mouton.
- Yamada, Yoshiko. 2010. A preliminary study of language contacts around Uilta in Sakhalin. *Hoppō jinbun kenkyū* 3. 59–75.
- Yamada, Yoshiko. 2016. Gishikutauda (Marīya Miheewa) no shōgai: Uiruta-go kita hōgen tekisuto. *Hoppō gengo kenkyū* 6. 179–201.
- Yamakoshi, Yasuhiro. 2006. Shinehen buriyāto-go tekisuto: Nichijō kaiwa o daizai ni shita kihon bunrei-shū. In Toshiro Tsumagari (ed.), *Languages of the North Pacific Rim*, vol. 13, 139–180. Sapporo: Graduate School of Letters, Hokkaidō University.
- Yamakoshi, Yasuhiro. 2007a. Hamuniga mongurugo tekisuto: Nichijō kaiwa o daizai ni shita kihon bunrei-shū. In Toshiro Tsumagari (ed.), *Languages of the North Pacific Rim*, 119–157. Sapporo: Graduate School of Letters, Hokkaidō University. vol. 14.
- Yamakoshi, Yasuhiro. 2007b. Shinehen buriyatogo no clitic. *Asian and African Languages and Linguistics* 2. 1–15.
- Yamakoshi, Yasuhiro. 2011a. Shineken Buryat. In Yasuhiro Yamakoshi (ed.), *Grammatical sketches from the field*, 137–177. Tōkyō: Research Institute for Languages, Cultures of Asia & Africa.
- Yamakoshi, Yasuhiro. 2011b. Three folktales in Shineken Buryat. *Asian and African Languages and Linguistics* 6. 109–136.
- Yamakoshi, Yasuhiro. 2015. Mongoru-go horuchin hōgen tekisuto: Nichijō kaiwa o daizai ni shita kihon bunrei-shū. *Hoppō gengo kenkyū* 5. 281–317.
- Yang Yonglong. 2014. Qinghai gangouhua fushu biaoji “[μ]” de leixing tezheng ji lishi bijiao. *Lishi yuyanxue yanjiu* 2. 239–254.
- Yeon, Jaehoon. 2012. Korean dialects: A general survey. In Nicolas Tranter (ed.), *The languages of Japan and Korea* (Routledge Language Family Series), 461–509. London: Routledge.
- Yeon, Jaehoon & Lucien Brown. 2011. *Korean. A comprehensive grammar* (Routledge Comprehensive Grammars). London: Routledge.
- Yibulaheimai A. 1985. Gansu jingnei tangwanghua jilüe. *Minzu yuwen* 6. 33–47.
- Yoon, Kyung-Eun. 2010. Questions and responses in Korean conversation. *Journal of Pragmatics* 42. 2782–2798.
- Yoshida, Yukata. 2009. Sogdian. In Gernot Windfuhr (ed.), *The Iranian languages* (Routledge Language Family Series), 279–335. London: Routledge.
- Yoshioka, Noboru. 2012. *A reference grammar of Eastern Burushaski*. Tōkyō: Tōkyō University of Foreign Studies. (Doctoral dissertation.)
- Young, Robert W. & William Morgan. 1987. *The Navajo language. A grammar and colloquial dictionary*. 2nd edn. Albuquerque: University of New Mexico Press.
- Young, S. 2006. Ukrainian. In Keith Brown (ed.), *Encyclopedia of language & linguistics*, 2nd edn., 212. Oxford: Elsevier.

## References

- Yu Wonsoo, Kwon Jae-il, Choi Moon-Jeong, Shin Yong-kwon, Bayarmend Borjigin & Luvsandorj[iin] Bold. 2008. *A study of the Tacheng dialect of the Dagur language* (Altaic Languages Series 2). Seoul: Seoul National University Press.
- Yunusbayev, Bayazit, Mait Metspalu, Ene Metspalu, Albert Valeev, Sergei Litvinov, Ruslan Valiev, Vita Akhmetova, Elena Balanovska, Oleg Balanovsky, Shahlo Turdikulova, Dilbar Dalimova, Pagbajabyn Nymadawa, Ardeshir Bahmanimehr, Hovhannes Sahakyan, Kristiina Tambets, Sardana A. Fedorova, Nikolay Barashkov, Irina Khidiyatova, Evelin Mihailov, Rita Khusainova, Larisa Damba, Miroslava Derenko, Boris Malyarchuk, Ludmila P. Osipova, Mikhail I. Voevoda, Levon Yepiskoposyan, Toomas Kivisild, Elza Khusnutdinova & Richard Villems. 2015. The genetic legacy of the expansion of Turkic-speaking nomads across Eurasia. *PLoS Genet* 11(4). 1–24.
- Zeitoun, Elisabeth. 2007. *A grammar of Mantauran Rukai* (Language and linguistics monograph series A4.2). Taipei: Academia Sinica.
- Zgusta, Richard. 2015. *The peoples of Northeast Asia through time. Precolonial ethnic and cultural processes along the coast between Hokkaido and the Bering Strait*. Leiden: Brill.
- Zhang Chengzai. 1980. Xining fangyan jilüe. *Fangyan* 4. 282–302.
- Zhang Dingjing. 1991. Hasakeyu yiwenju. *Xinjiang daxue xuebao* 61. 98–105.
- Zhang Yanchang, Li Bing & Zhang Xi. 1989. *The Qroqen language*. Jilin: Jilin daxue chubanshe.
- Zhang Yanchang, Zhang Xi & Dai Shuyan. 1989. *The Hezhen language*. Jilin: Jilin daxue chubanshe.
- Zhang, Paiyu. 2013. *The Kilen language of Manchuria*. Hong Kong: University of Hongkong. (Doctoral dissertation.)
- Zhao Jie. 1989. *Xiandai manyu yanjiu*. Peking: Minzu chubanshe.
- Zhao Xi. 1982. Chaoxianyu gaikuang. *Minzu yuwen* 3. 63–80.
- Zhao Xiangru & Aximu. 1981. Ainuyu de shuci. jianlun ainuyude xingzhi. *Minzu yuwen* 2. 44–48.
- Zhao Xiangru & Aximu. 1985. Xinjiang yilidiqu tu'erkehua de tedian. *Minzu yuwen* 5. 26–35.
- Zhao Xiangru & Aximu. 2011. *Ainuyu yanjiu*. Peking: Minzu chubanshe.
- Zhao Xiangru & Reinhard F. Hahn. 1989. The Ili Turki people and their language. *Central Asiatic Journal* 33. 260–289.
- Zhao Yong-Bin, Zhang Ye, Zhang Quan-Chao, Li Hong-Jie, Cui Ying-Qiu, Xu Zhi, Jin Li, Zhou Hui & Zhu Hong. 2014. Ancient DNA reveals that the genetic structure of the northern Han Chinese was shaped prior to 3,000 years ago. *PlosOne* 10(5). 1–17.
- Zhaonasiu. 1981a. *Dongbu yuguyu jianzhi*. Peking: Minzu chubanshe.
- Zhaonasiu. 1981b. *Tuzuyu jianzhi*. Peking: Minzu chubanshe.
- Zhong Jinwen. 2007. *Gansu diqu teyou minzu yuyan wenhua de quyuan tezheng*. Peking: Zhongyang minzu daxue chubanshe.
- Zhong Suchun. 1982. *Dawo'eryu jianzhi*. Peking: Minzu chubanshe.
- Zhu Yongzhong, Chaolu Wu, Keith W. Slater & Kevin Stuart. 1997. Gangou Chinese dialect. A comparative study of a strongly Altaicized Chinese dialect and its Mongolic neighbor. *Anthropos* 92(4/6). 433–450.

- Zhukova, A. N. 1997. Korjaskij jazyk. In T. Ju. Zhdanova, N. V. Gorova & O. I. Romanova (eds.), *Jazyki mira. paleoaziatskie jazyki*, 39–53. Moscow: Indrik.
- Zikmundová, Veronika. 2013. *Spoken Sibe. Morphology of the inflected parts of speech*. Prague: Karolinum.



# Name index

- Aalto, Pentti, 294, 323  
Abdurehim, Esmael, 341, 343  
Abondolo, Daniel, 33  
Acuo Yixiweisa, 52  
Adams, Douglas Q., 6, 150, 153–155, 158,  
159, 164, 165, 405  
Aikhenvald, Alexandra Y., 37, 269  
Aikio, Ante, 35  
Aixinjueluo Yingsheng, 307, 311  
Åkerman, Vesa., 238, 239, 242  
Alexander, Richard, 1  
Alimujiang Xiren, 163  
Allentoft, Morten E., 22, 437  
Alonso de la Fuente, José Andrés, 302,  
330, 416  
Amha, Azeb, 12, 75, 76  
An Jun, 299, 301, 327  
Anderson, Gregory D. S., 21, 44, 48–  
50, 113, 120, 345–348, 350–352,  
355, 360–362, 377, 386, 435  
Andvik, Erik E., 67  
Anthony, David W., 6, 18, 22, 23, 48, 437  
Aoi, Hayato, 181, 190  
Arakaki, Tomoko, 177, 178  
Araujo, Gabriel Antunes de, 62  
Arnheim, Rudolf, 60, 93  
Asai, Tōru, 111–113  
Aston, W. G., 189, 192, 194, 197  
Atknine, Victor, 284, 288, 320  
Audova, Iris, 13, 136  
Austerlitz, Robert, 116  
Avrorin, Valentin A., 299, 324, 325  
Axelsen, Jacob Bock, 4, 436, 438  
Axelsson, Karin, 55, 61, 69, 70  
Aximu, 343, 344, 361  
Baek, Sangyub, 312  
Bai Ping, 159, 160  
Baitchura, Uzbek, 301  
Bar, Moshe, 94, 412  
Baranesa, Adrien, 96, 101  
Barsalou, Lawrence W, 90–92, 94  
Bashir, Elena, 163  
Baskakov, N. S., 350, 360, 362  
Batchelor, John, 106, 111  
Baxter, William H., 28, 29, 46, 256, 274  
Beckwith, Christopher I., 25  
Bellwood, Peter, 10, 17  
Bencini, Giulia, 55, 73  
Bentley, John R., 19, 165, 190, 191  
Benzing, Johannes, 30, 230, 287, 289,  
290, 293, 312, 314–318, 320,  
322, 330  
Berge, Anna, 21, 22, 128  
Bergen, Benjamin, 91  
Berghäll, Liisa, 67, 68  
Bergsland, Knut, 128–130, 137, 139, 405  
Berlyne, Daniel E., 5, 95–97, 439  
Bhat, D. N. S., 55–57  
Bickel, Balthasar, 1, 2, 6, 85, 394, 435, 438  
Bickerton, Derek, 6, 7, 87, 417  
Bielmeier, Roland, 162  
Bilaniuk, Laada, 24  
Birjukovich, R. M., 351, 360  
Birtalan, Ágnes, 231, 250  
Bisang, Walter, 40, 41, 47  
Bläsing, Uwe, 231, 250  
Blench, Roger, 18, 28  
Blust, Robert, 191, 199  
Bobaljik, Jonathan D., 123  
Boeschoten, Hendrik, 32, 33, 344, 345,  
353, 354, 356, 359  
Bogoras, Waldemar, 126, 128

*Name index*

- Boldyrev, Boris V., 299, 324, 325  
Bolinger, Dwight, 54, 101  
Bowern, Claire, 55  
Bowern, Claire L., 58  
Braune, Wilhelm, 64, 140  
Brosig, Benjamin, 229  
Brown, Lucien, 200–203, 208  
Bugaeva, Anna, 19, 104–107, 110, 111  
Buhe, 234  
Bulatova, Nadezhda, 288, 318–320  
Busenitz, Robert L., 414
- Cable, Seth, 3, 56, 384, 385  
Campbell, Lyle, 39, 437  
Carling, Gerd, 150, 164  
Castrén, M. Alexander, 248, 318, 347,  
355, 361, 362, 369–373, 375,  
376, 380, 382, 383  
Chae, Heekyung, 115, 120  
Chaghanhada, 228–230  
Chaoke D. O., 291, 293, 294, 308, 322,  
323, 327  
Chaolu Wu, 222, 223, 232–234, 236, 246,  
252, 253  
Chard, Chester S., 10, 435  
Chen Litong, 52  
Chen Zhaojun, 241  
Chen Zongzhen, 335, 336, 349, 358, 359  
Cheng Mingyuan, 305–307  
Cheng, Andrew, 214  
Chien Yuehchen, 165  
Chirkova, Katia, 29, 30, 283  
Chisholm, William S., 54  
Cincius, Vera I., 312, 318, 319, 321  
Clark, Larry V., 354  
Clark, Marybeth, 61, 65, 260, 436  
Comrie, Bernard, 5, 40, 41, 43, 44, 49, 50,  
128, 146, 147, 377, 435, 437  
Cotrozzi, Stefano, 288, 320  
Couper-Kuhlen, Elizabeth, 63  
Creissels, Denis, 83  
Cribbs, Robert, 10, 11  
Croft, William, 71  
Cubberley, Paul, 15, 143, 145, 159, 160
- Cysouw, Michael, 5, 55, 76, 78, 80–82, 86,  
87, 111, 153, 157, 245, 407
- Dahl, Östen, 1, 40, 47, 51  
Danielsen, Niels, 54  
Davis, Christopher, 185, 187, 189, 191, 192,  
398  
De Busser, Rik, 4, 438  
De Reuse, Willem J., 24, 384, 385  
De Roerich, Georges, 269, 270  
De Sousa, Hilário, 43  
Dediu, Dan, 10, 438  
DeGiorgio, Michael, 17  
DeLancey, Scott, 29, 270, 281  
Delbrück, B., 140  
Denwood, Philip, 269  
Derksen, Rick, 158, 159  
Dewey, John, 95, 96  
Dexi, Zhu, 272  
Di Cosmo, Nicola, 306  
Dienst, Stefan, 75  
Diessel, Holger, 6, 53, 55, 76, 77, 80, 81,  
86, 89, 394  
Dik, Simon C., 59, 60  
Dillon, James T., 90, 412  
Ding Danqing, 247  
Ding, Picus Sizhi, 272  
Dingemanse, Mark, 37, 55  
Dixon, Robert M. W., 37, 53–55, 57–59,  
64, 70, 73, 76, 81, 89, 194, 240,  
241  
Doerfer, Gerhard, 32, 44, 46, 50, 51, 284,  
290, 293, 312, 315, 320, 323,  
356, 415  
Donidze, G. I., 349, 350, 362  
Doornenbal, Marius A., 72  
Dpal-ldan-bkra-shis, 253, 254  
Dryer, Matthew S., 4, 14, 41–43, 54, 55,  
61, 63, 141, 385, 395, 396, 436,  
438  
Duggan, Ana T., 31  
Duncker, Karl, 95  
Dunkel, George E., 153, 154  
Dunn, Michael, 122, 125–127

- Dwyer, Arienne M., 52, 264, 265, 280
- Ebata, Fuyuki, 351, 352
- Eberhard, Wolfram, 43
- Ebihara, Shiho, 15, 255, 268, 269, 281–283
- Ebina, Daisuke, 75
- Emmerick, Ronald E., 23, 24, 148, 163
- Enfield, Nicholas J., 2, 3, 40, 41, 43, 44, 53, 90, 435–438
- Enhebatu, Merden., 309, 327
- Enrico, John, 384
- Epps, Patience, 37, 38
- Erdal, Marcel, 32, 350, 353, 357, 362
- Evans, Nicholas, 1, 435, 437
- Evans, Vyvyan, 437, 438
- Everett, Caleb, 4, 438
- Everett, D. L., 59, 77
- Faehndrich, Burgel R. M., 237, 238, 253, 254
- Field, Kenneth L., 237, 253
- Fill, Alwin, 4, 90, 97, 99, 438
- Flegontov, Pavel, 17, 34, 437
- Foley, William A., 435
- Forsyth, James, 24, 35
- Fortescue, Michael, 16–18, 20–22, 46, 49, 113, 114, 117–121, 124, 125, 128, 137, 138, 405
- Fortson, Benjamin W., 23, 24, 46, 140, 152, 153
- Fountain, Amy, 385
- Frellesvig, Bjarke, 191
- Fried, Robert Wayne, 232–234, 244, 252
- Funk, Dimitrij A., 303
- Gallese, Vittorio, 91, 100
- Gao Erqiang, 24, 149, 150, 163, 164
- Gáspár, Csaba, 249
- Geiger, Wilhelm, 162
- Geng Shimin, 23, 24, 336, 338, 339, 359, 360
- Georg, Stefan, 20, 30, 46, 50, 123–126, 230, 238, 242, 247, 284, 377, 381, 382
- Gibson, Eleanor J., 96, 439
- Gibson, James J., 4, 90, 96–99, 101, 439
- Girfanova, Albina H., 297, 312
- Glenberg, Arthur M., 98
- Göksel, Asli, 333, 334, 354
- Golden, Peter B., 32
- Golovko, Evgenij V., 128, 130
- Gong Hwang-Cherng, 29, 255, 270, 271, 283
- Gorelova, Liliya M., 307, 312, 328
- Graczyk, Randolph, 38
- Graesser, Arthur C., 90
- Greenberg, Joseph H., 6, 13, 405, 406
- Grenoble, Lenore, 318, 319
- Grube, Wilhelm, 116–120, 318
- Gruzdeva, Ekaterina, 20, 113–116, 118–120
- Gusev, Valentin, 20, 52, 365, 366
- Haak, Wolfgang, 22, 437
- Hackstein, Olav, 55, 63, 73, 76, 86, 140, 150, 151, 153, 155
- Hagège, Claude, 55, 64
- Hahn, Reinhard F., 331, 344, 361
- Hai Feng, 280
- Hajdú, Péter, 363, 373
- Häkkinen, Jaakko, 18, 35
- Haller, Felix, 269, 282, 283
- Hammarström, Harald, 15, 44, 140
- Han Youfeng, 291, 322
- Harrison, K. David, 214, 345–347, 350–352, 355, 360–362
- Hasegawa, Toshikazu, 18, 19
- Hasegawa, Yoko, 15, 25, 165, 167, 169, 170, 172, 173, 194
- Hashimoto, Montaro J., 43
- Hasibate'er, 321
- Haspelmath, Martin, 14, 55, 66, 68
- Hauer, Erich, 306, 328, 330
- Hayashi, Makoto, 13, 172
- Hayashi, Yuka, 181

*Name index*

- Heggarty, Paul, 7, 11, 22, 40, 43, 437  
Heidermanns, Frank, 64, 140  
Heine, Bernd, 55, 72, 78–81  
Helimski, Eugen, 48, 366, 369, 370, 372, 374, 375  
Hellenthal, Anne-Christie, 62, 75, 76, 82  
Hengeveld, Kees, 55, 77–79  
Henrich, Joseph, 5  
Hinds, John, 167–169, 172  
Hirofumi, Matsumura, 19, 25  
Hodges, Bert H., 99, 439  
Hoeks, John C. J., 97  
Holton, Gary, 34  
Hölzl, Andreas, 1, 2, 12, 30, 37, 39, 51, 54–56, 58–61, 68, 70–72, 75, 76, 78, 79, 81–83, 86, 93, 198, 199, 229, 259, 284, 286, 287, 292, 316, 318, 328, 330, 365, 398, 399, 415, 416, 437, 438  
Hong Shi, 437  
Hoymann, Gertie, 65, 67  
Hu Zengyi, 291, 292, 294, 322  
Hu Zhenhua, 331, 339, 340, 348, 359, 362  
Huang Chenglong, 75, 272, 283, 417  
Huang, Lillian M., vii, viii, 6, 13, 55, 65, 66, 68, 94, 199, 200, 436  
Huddleston, Rodney, 53  
Hugjiltu, Wu, 233, 252  
Hyman, Larry M., 64, 438  
  
Idiatov, Dmitry, 37, 54, 55, 73, 81, 83, 86, 198, 319  
Iggesen, Oliver A., 42  
Ikegami, Jirō, 30, 39, 96, 120, 284, 296, 303, 304, 326  
Iksop, Lee, 208  
Imart, Guy, 331, 348, 362  
Ivanovskij, A. O., 294  
Izuyama, Atsuko, 165, 185–187, 190, 196  
  
Jacobs, Neil G., 24, 142, 143, 155, 156  
Jacobson, Steven, 133–137, 139  
Jacques, Guillaume, 271, 272, 282  
  
Janhunen, Juha, viii, 8, 10, 11, 15, 16, 18, 20, 25–27, 30–33, 35, 40, 44, 46, 48–52, 113, 217, 219–222, 224–226, 229–231, 244, 245, 248, 249, 265, 266, 269, 281, 284, 289, 312, 314–316, 319, 321, 331, 363, 364, 371, 373, 375, 376, 405, 414–416  
Järvilehto, Timo, 2, 90, 91, 438, 439  
Jedig, Hugo H., 141, 142, 144, 155  
Jeon, Hae-Sung, 210  
Jiang Li, 66, 271  
Jinam, Timothy, 19  
Johanson, Lars, 15, 32, 33  
Joki, A. J., 369, 376  
Jones, Eppie R., 22, 437  
  
Kałużyński, Stanisław, 308, 309, 323  
Kämpfe, Hans-Rainer, vii, 126, 127  
Kane, Daniel, 284  
Kang, Min Jeong, 97  
Kara, Dávid Somfai, 339  
Karlsson, Anastasia Mukhanova, 227  
Kasparov, Aleksey K., 48  
Katz, Dovid, 155  
Kazama, Shinjirō, 79, 295, 296, 302, 312, 315, 318, 323, 326, 416  
Kern, B., 59, 77  
Kerslake, Celia, 333, 334, 354  
Khabtagaeva, Bayarma, 51  
Khalilova, Zaira, 67  
Khasanova, Marina, 295, 296  
Kiaer, Jieun, 200, 207, 208, 214  
Kiefer, Ferenc, 57  
Kim, Deborah, 164  
Kim, Juwon, 309, 329  
Kim, Ronald I., 165  
Kim, Stephen S., 217, 236, 237, 251, 253  
Kim-Renaud, Young-Key, 202, 205  
Kimball, Geoffrey D., 64  
King, Ross J., 200, 201, 210, 214, 215  
Kirchner, Mark, 338, 339  
Kiyose, Gisaburo N., 284, 328  
Kizu, Mika, 167, 170

- Klamer, Marian, 94  
 Knüppel, Michael, 44, 290, 320  
 Ko, Dongho, 302  
 Kobayashi, Masato, 73, 406  
 Köhler, Bernhard, 12, 55, 62, 75, 76  
 Kokuritsu Kokugo Kenkyūjo, 170, 195  
 Koloskova, Yulia, 181  
 König, Ekkehard, 54, 55, 65  
 Kortt, I. R., 372  
 Kotorova, Elizaveta, 378, 379  
 Kraaijenbrink, Thirsa, 40  
 Krasovitsky, Alexander, 148  
 Kratochvíl, František, 82, 412  
 Krause, Wolfgang, 165  
 Krauss, Michael E., 384, 385  
 Krechevsky, Isadore, 95  
 Kroonen, Guus, 155  
 Künnap, Ago, 369, 371, 372, 376  
 Kupchik, John E., 25, 168–171, 190, 192,  
 194, 330  
 Kurpaska, Maria, 41, 255  
 Kuteva, Tania, 72  
  
 Ladstätter, Otto, 331  
 Laland, Kevin N., 2, 90, 96, 438  
 Landmann, Angelika, 333, 354, 356, 359  
 Langacker, Ronald W., 40, 44, 60, 77, 82,  
 91–93, 95, 438  
 LaPolla, Randy, 28, 75, 100, 267, 272, 283,  
 417  
 Lau, Tyler, 185, 187, 189, 191, 192, 398  
 Lawrence, Wayne P., 165, 183–185  
 Lbova, Ludmila, 10  
 Leben, William R., 64, 438  
 Lee-Smith, Mei W., 264, 265, 267, 331,  
 343, 361  
 Lehmann, Christian, 140  
 Levinson, Stephen C., 1, 4, 5, 10, 53, 55,  
 62, 90, 92, 93, 99, 398, 435, 437  
 Lewin, Kurt, 2, 4, 97, 99  
 Li Bing, 292, 293, 322, 327  
 Li Fengxiang, 292  
 Li Linjing, 315, 326  
 Li Yong-Söng, 351, 352, 360, 362  
  
 Li Zengxiang, 336, 338, 339, 359, 360  
 Li, Charles N., viii, 257, 260  
 Li, Fengxiang, 44, 51, 284  
 Lichtenberk, Frantisek, 55  
 Lie, Hiu, 284, 323  
 Liljegren, Henrik, 66, 73, 150, 151  
 Lin Lianyun, 334, 335, 358  
 Lindström, Eva, 55  
 Ling Chunsheng, 326, 327  
 Liu Liji, 261, 279  
 Liu Zhaoxiong, 236, 237, 252  
 Liuzhaoxiong, 234  
 Llamas, Bastien, 17, 437  
 Loewenstein, George, 4, 93, 96, 97, 439  
 Lopatin, Ivan A., 324  
 Luo Tianhua, 13, 55, 260  
  
 Ma Guoliang, 252  
 Ma Quanlin, 358  
 Ma, Jing-heng Sheng, 276, 277  
 Mackenzie, J. Lachlan, 6, 55, 77–79, 81,  
 85  
 Maddieson, Ian, 45–47  
 Majewicz, Alfred F., 110, 120, 326  
 Makelaike Yumai'erbai, 340  
 Malchukov, Andrej L., 284, 289, 290, 312  
 Mallory, James P., 6, 23, 150, 153–155, 158,  
 159, 164, 165, 405  
 Manrubia, Susanna, 4, 436, 438  
 Manzelli, Gianguido, 40  
 Martin, Samuel E., 195, 199, 246, 247, 249  
 Maslova, Elena, 18, 35, 386–394  
 Matayoshi, Satomi, 180, 183, 184, 197  
 Matić, Dejan, 35, 50, 290, 390, 391  
 Matsumori, Akiko, 170  
 Matthews, Stephen, 29, 41  
 Mattissen, Johanna, 117–120  
 Mauri, Caterina, 55  
 Mawkanuli, Talant, 347  
 McWhorter, John, 7, 38, 43, 87, 415, 416  
 Meng Shuxian, 291, 322  
 Menovshchikov, G. A., 133, 135  
 Menz, Astrid, 351, 363  
 Meyer, Michel, 54

*Name index*

- Mi Haili, 359  
Miestamo, Matti, 7, 13, 33, 55, 57, 58, 61,  
63, 64, 141, 364, 369, 370  
Mikola, Tibor, 368, 371  
Mithun, Marianne, 61, 70  
Miyake, Marc, 46  
Miyaoka, Osahito, vii, 82, 94, 131–137,  
139  
Miyara, Shinsho, viii, 176, 178, 179, 191,  
195, 199  
Moravcsik, Edith A., 54  
Moreno-Mayar, J. Víctor, 17  
Morgan, William, 384, 385  
Mostaert, Antoine, 230, 251  
Mu Yejun, 51, 284, 286, 311, 314, 316, 317,  
328, 329, 415  
Muhamedowa, Raihan, 15, 336–339, 342,  
354  
Mühlhäusler, Peter, 2  
Mus, Nikolett, 367, 371, 373, 375  
Mushin, Ilana, 55, 80, 82, 88  
Muysken, Pieter, 6, 7, 55, 81, 82, 87, 417  
  
Nagano, Yasuhiko, 282  
Nagano-Madsen, Yasuko, 179  
Nagasaki, Iku, 351, 386, 387, 389, 391–  
394  
Nagayama, Yukari, 121, 125–127  
Nakagawa, Hiroshi, 103  
Nakanome, Akira, 39, 304  
Nam, Pung-hyun, 26, 212, 213  
Napoli, Mateus Froes, 236, 237, 244, 253  
Narangoa, Li, 10, 11  
Nau, Nicole, 55, 85, 88  
NDSSLD, Neimeng dongbei  
shaoshuminzu shehui lishi  
diaochazu, 299, 300, 326  
Nedjalkov, Igor, 93, 286–288, 312, 321,  
322  
Nedjalkov, Vladimir, 123  
Nedjalkov, Vladimir P., 114–119  
Nefedov, Andrey, 378, 379  
Nettle, Daniel, 1, 436  
Nevskaja, Irina, 349, 350, 362  
  
Nichols, Johanna, 1, 6, 7, 11, 37, 40, 42, 45,  
47, 48, 85, 394, 405, 406, 435–  
437  
Nieuweboer, Rogier, 24, 143, 144, 155  
Niinaga, Yuto, 174, 175, 195, 196, 198, 199  
Nikolaeva, Irina A., 46, 94, 296–298,  
312, 324, 325, 367–369, 373–  
375, 386–388, 392, 393, 405  
NINJAL, Kokuritsu kokugo kenkyūsho,  
104–106, 110  
Norman, Jerry, vii, 28, 308, 311, 317, 329,  
330  
Novák, Lubomír, 40  
Nugteren, Hans, 232, 251  
Nuyts, Jan, 90  
  
OCLS, Okinawa Center of Language  
Study, 195, 196, 199  
Odling-Smee, John, 2, 3, 90, 96, 438  
Okuda, Osami, 103  
Olawsky, Knut J., 58  
Ömörüyi, Thomas O., 67, 73  
Operstein, Natalie, 16, 37, 87, 328, 415,  
416  
Otaina, Galina A., 114–119  
Oxenham, Marc, 19, 25  
O'Connor, Loretta, 55  
  
Pakendorf, Brigitte, 48–50  
Parker, Steve, 37  
Parpola, Asko, 33  
Pellard, Thomas, 25, 46, 165, 181, 182, 191,  
196, 198  
Peng Qiu, 170, 197–199  
Perls, Frederick S., 97  
Perry, John R., 149, 163  
Peterson, David A., 6, 42, 405  
Peterson, John, 89  
Pevnov, Alexander M., 18, 30, 39, 120,  
295, 296  
Peyraube, Alain, 78, 275  
Peyrot, Michaël, 164  
Pispane, Peter S., 35  
Pitulko, Vladimir V., 10, 48

- Poppe, Nicholas, 244, 323, 336, 359  
 Post, Mark W., 18, 28  
 Press, Ian, 146, 160, 161  
 Prins, Marielle., 271  
 Pugh, Stefan M., 146, 160, 161  
 Pulleyblank, Edwin G., 256, 274, 275, 277  
  
 Ragagnin, Elisabetta, viii, 227, 346, 347, 352, 356, 360  
 Raghavan, Maanasa, 17, 437  
 Rajasingh, V. R., 53, 76, 436  
 Ramsey, Robert S., 41, 208  
 Rassadin, V. I, 361  
 Ratliff, Martha, 41  
 Rawski, Evelyn S., 10, 11  
 Rédei, Károly, 35  
 Refsing, Kirsten., 107, 110  
 Reich, David, 10, 17, 21, 437  
 Reio, Thomas G, 5  
 Renfrew, Colin, 7, 11, 40, 43, 437  
 Renn, Jürgen, 2  
 Rialland, Annie, 55, 63  
 Rice, Karen, 57, 384, 385  
 Rimsky-Korsakoff, Svetlana, 263, 280  
 Ringe, Don, 6, 18, 22, 23, 48, 437  
 Robbeets, Martine, 44, 50  
 Røed, Knut H., 48  
 Róna-Tas, András, 46, 354  
 Roos, Martina Erica, 349, 357, 362  
 Rosol Christoph, Sara Nelson, 2  
 Ross, Claudia, 276, 277  
 Ross, John, 210  
 Ross, Lee, 92  
 Rossano, Federico, 53, 101  
 Rozycki, William, 316  
 Rybatzki, Volker, 27, 217, 221, 224, 231  
  
 Sadock, Jerrold M., 54, 60, 100, 128, 141  
 Safonova, Tatiana, 98  
 Sagart, Laurent, 28–30, 45, 46, 256, 274  
 Saltzman, Moira, 201, 207, 208, 214, 216  
 Sammallahti, Pekka, 46  
 San Roque, Lila, 100  
 Sanada Shinji, 165  
 Sanada, Shinji, 25  
 Sandman, Erika, 29, 52, 242, 255, 265–267, 414  
 Sanitt, Nigel, 13, 54, 90  
 Sankararaman, Sriram, 10  
 Sántha, István, 98  
 Schiefner, Anton, 290, 320, 387, 389, 390  
 Schmalz, Mark, 388, 391, 392, 394  
 Schmidt, Peter, 300, 302, 312, 318, 325  
 Schmitt, Rüdiger, 23  
 Schönig, Claus, 27, 33, 51, 347, 354, 356, 361, 405  
 Schröder, Dominik, 253  
 Schulze, Wolfgang, 4, 53, 55, 82, 86, 90, 95, 100, 324  
 Sean, Lee, 18, 19, 25, 200  
 Sechenbaatar, Borjigin, 217, 227, 249  
 Seebold, Elmar, 155, 158  
 Sekerina, Irina A, 128, 139  
 Sem, Lidija I., 326  
 Serafim, Leon A., 187  
 Shapiro, Roman, viii, 147, 162  
 Shaw, Robert B. S., 149, 340–342, 344  
 Shevelov, George Y., 145, 146, 160  
 Shi Feng, 279  
 Shibatani, Masayoshi, 103, 108, 110, 111, 165, 169  
 Shigeno, Hirumi, vii, 173  
 Shimoji, Michinori, 165, 182, 183, 187, 189, 191  
 Shimunek, Andrew, 27, 32, 51, 377  
 Shinzato, Rumiko, vii, 171, 187, 188  
 Shiraishi, Hidetoshi, 114, 117–119  
 Sicoli, Mark A., 34, 63, 436  
 Sieg, Emil, 150, 164  
 Siegl, Florian, 366, 372–374  
 Siegling, Wilhelm, 150, 164  
 Siemund, Peter, 54, 55, 57, 63, 65, 68, 88, 398  
 Simčenko, Ju. B., 372  
 Simon, Camille, 29, 52, 255  
 Sinha, Chris, 2, 90  
 Sinor, Denis, 8, 33

*Name index*

- Siqinchaoketu, 217, 234, 235, 251, 252  
Siska, Veronika, 25, 31  
Skoglund, Pontus, 437  
Skribnik, Elena, 49, 224, 248, 437  
Slater, Keith W., 52, 240–242, 254, 414  
Smith, Norval, 6, 7, 55, 81, 82, 87, 417  
Sohn, Ho-Min, 26, 200, 202–205, 207–217  
Song, Jae Jung, 15, 201–204, 206, 208, 214, 215  
Sotavalta, Arvo, 290, 320  
Spencer, Andrew, 126  
Stachowski, Marek, 351, 352, 354, 355, 363  
Steffensen, Sune Vork, 4, 90, 97, 99, 438  
Stern, Dieter, vii, 147, 148, 161  
Stibbe, Arran, 1  
Stilo, Donald L., 40  
Stivers, Tanya, 53, 55  
Stolz, Thomas, 84  
Street, John C, 221  
Sulamo, Dagnachew Degu, 70, 76  
Sumner, Francis B., 2  
Sun Chaofen, vii, 260, 265  
Sun Hongkai, 82, 88, 271, 273, 281, 282  
Sun Qinglin, 279  
Sun, Jackson T.-S., 255, 269–272, 281, 282  
Sunik, Orest, 323  
Sussex, Roland, 143, 145, 160  
Suzuki, Hiroyuki, 282  
Svantesson, Jan-Olof, 226, 227  
Swenson, Rod, 93, 99  
Szokolszky, Agnes, 90  
  
Taaffe, Robert N., 10  
Takahashi, Yasushige, 108, 109  
Takehiro, Sato, 19  
Takuichiro, Onishi, 170  
Tamura, Suzuko, 104–106, 110  
Tangiku, Itsuji, 114, 117–119  
Taylor, Archer, 96  
the APiCS Consortium, 55  
Thomas, Werner, 165  
  
Thomason, Sarah G., 37, 38  
Thompson, Hanne-Ruth, 62, 73, 151, 406  
Thompson, Sandra A., viii, 257, 260  
Tietze, Andreas, 331  
Tirkey, Bablu, 73, 406  
Tittel, Hans (transl.), 108, 109, 111  
Todaeva, Buljas Ch., 236, 237, 252  
Tokunaga, Akiko, 176, 181, 190, 193, 194, 196  
Tolskaya, Inna, 59, 298, 315, 324  
Tolskaya, Maria, 59, 94, 296–298, 312, 315, 324, 325  
Tomasello, Michael, 4, 79, 91–93, 98, 99  
Tomomi, Sato, 108  
Tooru, Hayasi, 331, 343, 347  
Toshikazu, Hasegawa, 18, 25  
Toshio, Ohori, 181  
Tournadre, Nicolas, 29, 100, 255, 267, 281  
Tranter, Nicolas, 25, 165, 167, 170  
Treis, Yvonne, 75, 76  
Trudgill, Peter, 5, 7, 87, 416  
Tsumagari, Toshiro, 15, 31, 35, 39, 49, 222, 246, 247, 293, 303, 304, 318, 323, 326  
Tuohuti Litifu, 15, 341–344, 359  
Turvey, Michael T., 2, 93, 99, 439  
  
Uemura, Yukio, 25, 177, 178  
Ultan, Russell, 54  
  
Vajda, Edward J., vii, 17, 18, 34, 49, 377, 379, 381–383  
Vakhtin, Nikolai, 128, 130, 133, 134  
Van Alsenoy, Lauren, 56  
Van den Berg, René, 414  
Van der Auwera, Johan, 55, 56, 73, 90  
Van der Lubbe, Gijs, 176, 181, 190, 193, 194, 196  
Van Driem, George, 28, 255  
Vandamme, Marc, 32, 33, 353, 356  
Veluppilai, Viveka, 37  
Voegelin, Carl F., 12, 436, 438  
Voegelin, Florence M., 12, 436, 438  
Volodin, Aleksandr P., 122, 124

- Volodin, Alexander P., vii, 20, 123–127  
Volodko, Natalia V., 35  
Von Möllendorff, Paul G., 306  
von Stumm, Sophie, 5  
Vovin, Alexander, 16, 18–20, 22, 25, 27,  
    30, 34, 46, 50, 51, 103, 109, 111,  
    168, 169, 171, 189–194, 196, 200,  
    216, 377, 405, 415
- Wade, Terence, 160  
Wagner-Nagy, Beáta, 369, 374, 376  
Wang Qingfeng, 310, 311, 329  
Watters, David E., 73, 80  
Weiers, Michael, 223  
Weiqiang, Hei, 279  
Welsch, Wolfgang, 2  
Werner, Heinrich, 377–383  
Whaley, Lindsay J., 44, 51, 284, 291  
Whitman, John, 26, 46, 51, 191  
Wiener, Gerald, 48  
Wilbur, Joshua, 64, 81, 88, 89  
Windfuhr, Gernot, 149, 163  
Winter, Werner, 23  
Witsen, Nicolaas, 10  
Wu Hongwei, 346, 361  
Wu, Fuxiang, 78, 275  
Wuge Shouping, 305–307  
Wurm, Stephen A., 265, 331, 343  
Wurmbrand, Susi, 123  
Wylie, Alexander, 305–307
- Xie Xiaoan, 263–265  
Xiren Kuerban, 163  
Xu, Dan, 267, 275, 280, 281  
Xuan Dewu, 201–205, 214
- Yamada, Masahiro, 186, 196  
Yamada, Yoshiko, 19, 20, 39, 52, 304  
Yamakoshi, Yasuhiro, 93, 220, 224, 228,  
    229, 247–249  
Yang Yonglong, 280, 281  
Yeon, Jaehoon, 200–202, 208, 209  
Yi Liqian, 335, 336, 358, 359  
Yibulaheimai A., 267
- Yongliang, Leng, 279  
Yoon, Kyung-Eun, 13, 205, 206, 214  
Yoshida, Yukata, 24, 148, 149, 162, 163  
Yoshioka, Noboru, 149, 381  
Young, Robert W., 384, 385  
Young, S., 15  
Yu Wonsoo, 223, 246  
Yunusbayev, Bayazit, 18, 331  
Yurn, Gyudong, 302
- Zeitoun, Elisabeth, 68  
Zgusta, Richard, 11  
Zhang Chengzai, 262, 279  
Zhang Dingjing, 13, 336–339  
Zhang Shumin, 263–265  
Zhang Xi, 300, 301  
Zhang Yanchang, 292, 293, 300, 301, 322,  
    327  
Zhang, Paiyu, 298, 300, 301, 315, 327, 416  
Zhao Jie, 310, 329  
Zhao Xi, 201, 203  
Zhao Xiangru, 343, 344, 361  
Zhao Yong-Bin, 43  
Zhaonasiu, 232, 241, 251  
Zhong Jinwen, 52, 280  
Zhong Suchun, 222, 246  
Zhu Yongzhong, 240, 267, 281  
Zhukova, A. N., 122  
Zikmundová, Veronika, 308, 329  
Zwicky, Arnold M., 54, 60, 141



# Language index

- Abui, 81, 94, 412  
Afro-Asiatic, 76  
Afroasiatic, 12, 44, 58, 64, 70, 75, 76  
Ainu, 15, 16, 19, 20, 52, 103–109, 111, 171,  
228, 400, 402, 404, 406, 411,  
419, 442, 445  
Ainu-Itelmen hybrid, 16, 38  
Ainuic, 15, 16, 18–21, 25, 45, 46, 49, 50,  
103, 109, 111, 169, 244, 354, 398,  
401, 405, 418  
Alchuka, v, 16, 51, 213, 284, 286, 311, 313–  
317, 327–330, 405, 413  
Aleut, 6, 16, 21, 22, 24, 38, 127–130, 132,  
136, 137, 140, 159, 397, 399, 402,  
405, 411, 413, 416, 419, 442, 445  
Altai Low German, 24, 141, 142, 144, *see*  
Plautdiitsch  
Altai Turkic, 332, 349, 350, 360, 420  
Alutor, 20, 120, 121, 124–127, 397, 400,  
419, 442, 445  
Amdo Tibetan, 15, 29, 33, 49, 100, 255,  
268–271, 273, 281, 283, 397,  
400, 403, 408, 409, 420, 443,  
446  
Amis, 65, 94  
Amuric, 15, 16, 18–21, 25, 31, 45, 46, 49,  
113, 117, 121, 124, 246, 303, 397,  
398  
Ancient Egyptian, 10  
Anong, 82, 88  
Apache, 384, 385  
Arabic, 16, 32  
Aramaic, 16, 24  
Arawakan, 55  
Arawan, 75  
Arghu, *see* Khalaj  
Arin, 34, 377  
Arman, 44, 285, 311, 314, 319, 320  
Assan, 34, 377  
Assanic, 377  
Atayal, 165, 170, 170<sup>13</sup>, 170<sup>14</sup>, 191, 199, 200  
Athabaskan, 17, 384  
Austroasiatic, 19, 41, 44, 65, 76, 89  
Austronesian, 10, 13, 15, 41, 65, 66, 68, 94,  
165, 191, 198, 200, 414  
Avestan, 152, 153  
Azeri, 46  
Baima, 29, 30, 255, 270–273, 281, 420,  
443, 446  
Bala, 16, 51, 213, 284, 286, 311, 313–315,  
327–330, 415  
Balaic, 286  
Balantak, 94, 414  
Baltic, 22, 80, 140  
Bantawa, 72, 73  
Barbacoan, 100  
Bardi, 58, 71, 73  
Bashkir, 46  
Basque, 44  
Belorussian, 141  
Bengali, 61, 62, 73, 151, 406  
Bonan, 217, 218, 232–236, 242–245, 250,  
252, 265, 402–404, 414, 415,  
419, 443, 446  
Breton, 154  
Burmese, 28  
Burushaski, 149, 381, 404  
Buryat, 15, 27, 31, 33, 49, 84, 217–221, 223,  
224, 242–247, 287, 293, 360,  
400, 404, 408, 409, 419, 443,  
446

*Language index*

- Chagatay, 32, 33, 332, 353, 354  
Chakhar Mongolian, 217, 227  
Chalkan, 94, 332, 350, 353, 355–357, 360, 413  
Chapacuran, 77  
Cha'palaa, 100  
Chinese, 4, 7, 10, 12, 13, 16, 19, 25, 28, 29, 43, 44, 52, 147, 162, 170, 198, 199, 199<sup>18</sup>, 200, 202, 218, 220, 222, 229<sup>21</sup>, 232–234, 236, 239, 241, 250, 251, 254, 255, 255<sup>22</sup>, 256, 259–261, 263, 264, 267, 272, 273, 275, 278, 280, 284, 292, 293, 298–301, 305, 306, 309, 311<sup>30</sup>, 331, 337, 339, 343, 352, 353, 396, 401, 403, 413–415, 417, 444, 447  
Chinese Pidgin Russian, 16, 38, 140, 147, 162  
Chukchi, 21, 22, 35, 45, 48, 120–127, 369, 400, 409, 419, 442, 445  
Chukotian, 121, 125  
Chukotko-Kamchatkan, 15, 18, 20, 21, 32, 35, 45, 46, 50, 113, 120, 121, 123–126, 128, 137, 290, 385, 387, 388  
Chulym, 332, 351, 353, 356, 357, 360, 362, 404, 420, 444, 447  
Chuvan, 386, 387  
Chuvash, 11, 32, 46, 94, 331, 332, 353–356  
Classical Tibetan, 29, 270  
Common Dialectal Chinese, 28  
Cone, 29, 255, 281, 282  
Crow, 38, 61  
Czech, 143  
Dagur, 27, 32, 52, 217, 218, 222, 223, 225, 228, 229, 242, 243, 245–247, 250, 291, 292, 307, 309, 312, 322, 323, 398, 404, 413, 419, 443, 446  
Daohua, 52  
Dardic, 66, 73, 149, 150, 404  
Darkhat, 227, 229  
Dene-Yeniseian, 10, 15, 17, 34, 383  
Djabugay, 85  
Dolgan, 33, 38, 244, 332, 334, 335, 352, 353, 355–357, 360, 387, 397, 400, 401, 404, 420, 444, 447  
Dravidian, 73, 406  
Dukhan, 332, 334, 346, 347, 352, 353, 356, 357, 360, 380<sup>37</sup>, 404  
Dunan, 166, 185–188, 190, 193, 195  
Dungan, 255, 263  
Dutch, 16, 154, 156, 158  
Edo, 67, 73  
Enents, 364  
Enets, 33, 348, 363, 364, 366, 367, 369–371, 374, 375, 395, 400, 403, 420, 444, 447  
English, v, 6, 10, 12, 15, 16, 21, 22, 24–26, 53, 57, 58, 60, 61, 63, 64, 66, 69, 70, 74, 76, 78, 80–87, 91, 92, 140, 141, 141<sup>5</sup>, 142–144, 146, 148, 152–154, 156–158, 172, 181<sup>15</sup>, 182, 254, 257, 319, 369, 388, 397, 407–411, 417  
Eskaleut, 16–18, 21, 34, 45–47, 94, 100, 128, 130<sup>3</sup>, 136, 140, 354, 419  
Eskimo, 128, 129, 137, 405  
Estonian, 33, 364  
Eurasianic, 13, 406  
Even, v, 7, 21, 30, 31, 35, 38, 44, 48, 49, 52, 84, 98, 124, 151, 165, 231, 284, 285, 289–291, 294, 295, 302, 312–320, 325, 380, 402, 404, 420, 443, 446  
Evenki, 2, 13, 20, 30–34, 38, 44, 48, 49, 52, 76, 86, 87, 93, 98, 130, 220, 224, 284–289, 291, 293–297, 302, 305, 312–315, 317–320, 322–324, 389, 402, 404, 408–411, 413, 417, 420, 443, 446  
Ewenic, 30, 31, 48, 49, 242, 284, 285, 294, 312, 314, 315, 317, 319, 322–324, 404, 413, 417  
Eyak, 384, 385

- Eynu, 16, 24, 38, 140, 331, 332, 340, 343, 353, 356, 360, 361
- Finnic, 363
- Finnish, 12, 40, 46, 63, 64, 66, 364
- Finno-Mordvinic, 363
- Finno-Permic, 363
- Finno-Saamic, 363
- Finno-Ugric, 33, 34, 40, 363
- Finno-Volgaic, 363
- French, 86, 181<sup>15</sup>
- Fuyu, 331, 332, 348, 353, 356, 360
- Gangou, 16, 38, 52, 255, 267, 273, 278, 415
- German, 12, 24, 37, 47, 69, 70, 73, 77–79, 84, 85, 87–89, 101, 140, 141, 141<sup>6</sup>, 142, 142<sup>7</sup>, 143, 144, 144<sup>8</sup>, 146, 150–154, 156–159, 182, 397, 400, 408–411
- Germanic, 7, 22–24, 64, 140, 141, 152, 153, 158, 364
- Gilyak, 20
- Gothic, 64, 140, 141, 402
- Govorka, 16, 38, 140
- Great Andamanese, 44
- Greek, 22
- gSerpa, 29, 255, 271, 272, 281
- Guiqiong, 66, 271, 283
- Hachijō, 25, 165, 166, 170, 188, 190, 191, 193, 195, 419
- Hadiyya, 70, 76
- Halkomelem, 64
- Hateruma, 166, 183, 184, 187, 188, 190, 193, 197, 198, 397, 419, 442, 445
- Hatoma, 166, 184–188, 190, 403, 419, 442, 445
- Hausa, 64
- Hawai'i Creole English, 37
- Hebrew, 16, 24
- Hezhen, 299, 300, 300<sup>26</sup>, 302, 326
- Hirara, 166, 181, 190, 191
- Hmong-Mien, 41
- Hungarian, 46, 364
- Ikema, 166, 176, 181, 187, 188, 190, 401, 419, 442, 445
- Ili Turki, 331, 332, 344, 353, 356, 360
- Indo-Aryan, 140
- Indo-European, 15, 16, 18, 22–24, 28, 33, 38, 44, 45, 47, 48, 63, 64, 66, 73, 121, 140, 150–153, 158, 163, 165, 375, 381, 385, 405, 419, 437
- Indo-Iranian, 22, 23, 62, 140
- Irabu, 166, 167, 181–183, 187–193, 401, 419, 442, 445
- Iranian, 23, 24, 33, 140, 148, 152, 159, 162, 163, 355, 413
- Ishigaki, 183, 185, 190
- Isoko, 65
- Italian, 84, 86, 87, 101, 154, 405, 406, 410
- Itelmen, 19–21, 121–125, 127, 128, 405, 419, 442, 445
- Japanese, v, 12, 13, 15, 16, 19, 20, 25, 29, 39, 43, 45, 52, 105, 126, 165, 165<sup>12</sup>, 166–170, 172–174, 179, 181, 187–200, 206, 215, 216, 221, 228, 266, 275<sup>24</sup>, 304<sup>28</sup>, 398, 400, 401, 403, 408, 409, 411, 419, 436, 442, 445
- Japhug, 272
- Japonic, 16, 18, 19, 25, 26, 39, 46, 47, 50, 51, 58, 165, 169, 183, 187–191, 193, 205, 213, 244, 283, 354, 391, 397, 398, 401, 413, 418, 419
- Jeju, 25, 200, 201, 207–209, 214–216, 395, 403, 419, 443, 446
- Jurchen A, 284, 286, 311
- Jurchen B, 284, 286, 311
- Jurchenic, 30, 213, 284–286, 294, 305, 311, 312, 315, 316, 325–328, 330, 404, 413, 415, 417
- Kalmyk, 26, 27, 94, 230, 231, 243, 245, 250
- Kamass, 33, 363, 364, 370, 375, 380, 404
- Kamchadal, *see* Itelmen

*Language index*

- Kangjia, 217, 218, 234–236, 243–245, 250,  
251, 253, 402, 404, 413, 415, 419,  
443, 446
- Kansai, 194
- Karagas, 347, 355–357
- Kazakh, 13, 15, 32, 332, 335–340, 342–  
344, 353, 354, 356, 358, 360,  
396–398, 401, 403, 420, 444,  
447
- Kenaboi, 44
- Kerek, 20, 21, 35, 121, 122, 124, 125, 400
- Ket, v, 15, 16, 34, 46, 47, 50, 94, 126, 369,  
377–382, 384, 385, 396, 397,  
402, 404, 409, 411, 417, 420,  
444, 447
- Khakas, v, 246, 332, 347, 348, 353, 356,  
360, 370, 405, 408, 409, 413,  
420, 444, 447
- Khalaj, 32, 332, 355, 356, 360
- Khalkha, v, 217, 224–228, 230, 232, 246,  
247, 397, 400, 401, 417, 419,  
443, 446
- Kham, 73
- Khamnigan Mongol, 32, 217–220, 224,  
242–245, 287, 288, 360, 400,  
404, 419, 443, 446
- Khantic, 40, 363
- Khanty, 40, 364
- Kharia, 89
- Khitano, 26, 27, 46, 50, 51
- Khitano-Mongolic, 27, 32, 49, 51
- Khoe, 67
- Khorchin, 31, 52, 93, 126, 217, 228–231,  
243, 247, 284, 292, 293, 308,  
398, 400, 404, 411, 417
- Khotanese, *see* Saka
- Khotong, 32
- Khwarshi, 66, 67
- Kilen, 16, 285, 298–300, 300<sup>26</sup>, 301, 306,  
312–315, 317, 319, 324, 326, 327,  
329, 330, 398, 401, 403, 404,  
413, 416, 417, 420, 443, 446
- Kili, 16, 285, 311, 314, 315, 323, 324, 327,  
413
- KiNubi, 87
- Kipchak, 32, 331, 332, 335
- Kipchak Uzbek, 332
- Koasati, 64
- Kohama, 190
- Koibal, 347, 356
- Kolyemal, 26, 201, 210, 215, 408, 409
- Komi-Zyrian, 364
- Korean, v, 13, 15, 16, 25, 26, 29, 46, 47,  
117, 168, 169, 192, 200–216, 220,  
225, 228, 266, 307, 322, 330,  
395, 396, 398, 399, 401, 403,  
411, 416, 419, 436, 443, 446
- Koreanic, 16, 18, 25, 26, 31, 46, 47, 49–51,  
167, 191, 200, 213, 283, 311, 396,  
398, 404, 413, 415, 416
- Koryak, 20, 121, 122, 124–128, 397, 400,  
419, 442, 445
- Kott, 34, 370, 377, 380–382, 404
- Kulina, 75
- Kuroshima, 190
- Kurux, 73, 406
- Kusunda, 80, 165
- Kyakala, 51, 286, 311
- Kyakhta Pidgin, 140
- Kyrgyz, 331, 332, 335, 339, 340, 343, 344,  
353, 356, 358, 360, 420, 444,  
447
- Latin, 16, 64, 87, 150, 152, 159
- Linxia Chinese, *see* Hezhou Chinese
- Lithuanian, 154
- Maltese, 44
- Manchu, 7, 27, 30, 31, 38, 43, 49–52,  
60, 76, 77, 83–86, 126, 213,  
215, 223, 228, 247, 284–286,  
297, 298, 300, 300<sup>26</sup>, 305–311,  
311<sup>30</sup>, 312–318, 320, 322, 324–  
330, 398, 402–404, 407–409,  
411, 413, 415–417
- Manchu-Tungusic, 30, *see* Tungusic
- Manchuic, 284, 286, 327

- Mandarin, v, 11, 13, 15, 29, 31, 32, 38, 41, 43, 46, 47, 49, 50, 60, 68, 70, 73, 79, 83, 147, 147<sup>10</sup>, 148, 198, 199, 224, 229, 234, 237, 247, 251, 254–263, 263<sup>23</sup>, 264–266, 270, 272–275, 278, 280, 307–310, 312, 335, 337, 397, 399–402, 404, 406–411, 413, 415, 417, 420, 436, 443, 446
- Mangghuer, 218, 234, 237, 240–243, 245, 250, 254, 402, 404, 413, 415, 419, 443, 446
- Mansi, 40, 364
- Mansic, 40, 363
- Mari, 46, 364
- Mariic, 363
- Mator, 33, 363, 369–371, 375
- Mauwake, 60, 67, 68
- Mednyj Aleut, *see* Copper Island Aleut
- Middle High German, 24
- Middle Korean, 211, 212, 215, 216, 398, 404, 406, 415
- Middle Mongol, 27, 50, 168, 221, 225, 227, 238, 244, 333, 397
- Miyara, 166, 185, 187–193, 195, 398, 399, 401, 419, 442, 445
- Moghol, 26, 27, 163<sup>11</sup>, 217, 218, 224, 242, 243, 245
- Mon, 65
- Mongghul, 218, 237–243, 245, 253, 254, 415, 419, 443, 446
- Mongolian, 15, 24, 27, 52, 93, 168, 217, 219, 221, 222, 224–232, 243–245, 247, 249, 250, 284, 291–295, 302, 308, 334, 347, 380<sup>37</sup>, 396, 404, 407–409, 411, 417
- Mongolic, v, 16, 18, 26, 27, 31–33, 35, 43, 46, 48–52, 93, 106, 119, 119<sup>2</sup>, 147<sup>9</sup>, 205, 213, 217–220, 222–224, 226, 227, 231, 232, 236, 240, 242, 244–246, 251, 255, 263, 265, 269, 281, 287, 288, 291, 293, 303, 308, 316, 319, 322, 331, 334, 347, 352, 355, 360, 398, 400–402, 404–406, 413, 415
- Mordvin, 364
- Mordvinic, 363
- Muskogean, 64
- Muöt, 76
- Na-Dene, 15–17, 34, 57, 383–385
- Nagahama, 190
- Nakachi, 166, 183
- Nanai, 6, 16, 30, 39, 79, 86, 246, 285, 295, 299, 300, 300<sup>26</sup>, 302, 305, 312–318, 324–326, 328, 405, 413, 416, 417, 420, 443, 446
- Nanaic, 30, 284, 285, 301, 302, 312, 315, 316, 319, 323, 326, 416, 417
- Naukan, 21, 129, 132, 134–137
- Naukanski, *see* Naukan
- Navajo, 384, 385
- Negidal, 20, 44, 52, 136, 231, 285, 289, 294–296, 302, 312–314, 317, 318, 320, 370, 392, 396, 403, 404, 420, 443, 446
- Nembe, 64
- Nenets, 33, 122, 348, 363, 364, 367–371, 373–375, 395, 397, 400, 403, 405, 408, 409, 420, 444, 447
- Nganasan, 16, 33, 35, 136, 148, 159, 161, 363–367, 370, 371, 374, 375, 395, 397, 400, 403, 413, 417, 420, 444, 447
- Niger-Congo, 15, 64, 65, 67, 73
- Nivkh, 15, 16, 19, 20, 31, 39, 48, 52, 113–117, 119, 120, 125, 304, 319, 325, 399, 402–404, 411, 413, 419, 442, 445
- Nuristani, 23, 140
- Nymylan, *see* Itelmen
- Nyulnyulan, 58, 71
- Ob-Ugric, 40
- Odul, *see* Yukaghir

*Language index*

- Ōgami, 166, 176, 181, 182, 187, 188, 190, 192, 193, 198, 228, 398, 400, 401, 403, 419, 442, 445
- Oghur, 32, 332, 354, 405
- Oghuz, 32, 331–334, 354
- Oirat, 15, 26, 217, 218, 230, 231, 243–245, 250, 400, 403, 419, 443, 446
- Okinawan, *see* Shuri
- Okinoerabu, 166, 176, 181, 183, 185, 186, 188, 190, 193, 194, 196, 403, 419, 442, 445
- Old Chinese, 10, 29, 46, 47, 256, 273–275, 410
- Old Japanese, 25, 39, 109, 166, 168–171, 179, 187, 188, 190–196, 214, 216, 330, 333, 397, 404
- Old Korean, 213
- Old Ryūkyūan, 169, 188, 190, 192, 216
- Old Tibetan, 29
- Ometo, 55
- Omok, 386
- Omotiic, 12, 75, 76
- Ongan, 44
- Ordos, 218, 230, 232, 243, 245, 247, 250
- Oroch, 285, 298, 299, 312–317, 324, 325, 413
- Oroqen, v, 27, 32, 44, 51, 52, 223, 225, 229, 247, 284, 285, 291–294, 298, 306, 312–314, 318, 320, 322–324, 327, 397, 401, 403, 404, 410, 413, 420, 443, 446
- Ostyak-Samoyed, *see* Selkup
- Paleo-Asiatic, 15
- Paleo-Eskimo, 17
- Paleo-Siberian, *see* Paleo-Asiatic
- Palula, 66, 73, 149–151
- Pama-Nyungan, 85
- Para-Ainuic, 103
- Para-Japonic, 25, 26, 165
- Para-Koreanic, 26
- Para-Mongolic, 26, 27, 31, 52
- Para-Yeniseic, 377
- Permic, 363
- Persian, 16, 24, 140, 149, 163, 163<sup>11</sup>, 331, 340, 343
- Pichis Ashéninca, 55, 87
- Plautdiitsch, 24, 141, 142, 144, 144<sup>8</sup>, 152–154, 156–158, 396, 397, 419, 442, 445
- Polish, 141, 154
- Portuguese, 16
- Prakrit, 16, 23
- pre-modern Japanese, 187, 189, 190, 192–194
- Prinmi, 272
- Proto-Indo-European, 152, 158
- Proto-Japonic, 189, 191
- Proto-Ainuic, 46, 47, 111
- Proto-Austronesian, 191
- Proto-Dene-Yeniseian, 34
- Proto-Eskaleut, 22, 46
- Proto-Eskimo, 138
- Proto-Indo-European, 6, 22, 85, 153, 373
- Proto-Japonic, 46, 169, 171, 188, 189, 192, 216
- Proto-Mongolic, 6, 26, 27, 219, 250, 254
- Proto-Ryūkyūan, 191, 192
- Proto-Samoyedic, 33, 46, 371
- Proto-Sino-Tibetan, *see* Proto-Trans-Himalayan
- Proto-Slavic, 159
- Proto-Tibetic, 281
- Proto-Tungusic, 30, 46, 51, 220, 221, 247, 287, 312, 314, 316, 317, 325, 333, 410
- Proto-Turkic, 6, 46, 354
- Proto-Uralic, 33, 46
- Proto-Yeniseic, 34, 377, 382, 383
- Proto-Yukaghiric, 6, 35, 46, 392
- Proto-Yupik-Sirenik, 138
- Pumpokol, 34, 377
- Qiang, 75, 100, 272, 283, 417
- Qiangic, 28–30, 66, 75, 255, 270–272, 281, 283
- Quechua, 74

- rGyalrong, 271, 272  
Rouran, 16, 50  
Rukai, 68  
Russian, v, 6, 12, 15, 16, 19–22, 24, 27, 29,  
32, 34, 35, 44, 49, 128, 130, 130<sup>3</sup>,  
137, 140, 145–148, 152, 153, 158–  
162, 247, 288, 298, 302, 303<sup>27</sup>,  
304<sup>28</sup>, 305, 350, 352, 360, 369,  
375, 378, 380, 382, 387, 396,  
397, 402, 404, 413, 416, 417, 419,  
442, 445  
Ryūkyūan, 16, 25, 165, 169–171, 173, 179,  
181, 184, 187, 190–197, 396, 398,  
401  
Saami, 64, 81, 88, 364  
Saamic, 363  
Sabanê, 61, 62  
Saisiyat, 65, 66  
Saka, 24, 148  
Salar, 265, 331, 332, 334, 335, 352–354,  
356, 358, 400, 403, 404, 414,  
420, 444, 447  
Salish, 64  
Samagir, 312  
Samar, 300, 312  
Samoyedic, 31, 33, 35, 45, 48, 128, 137,  
348, 363, 364, 366, 370, 371,  
373, 375, 396, 405, 435  
Sanskrit, 16, 23, 151  
Santa, 15, 217, 218, 234, 236, 237, 243–  
245, 250–252, 402, 404, 413,  
415, 417, 419, 443, 446  
Sanuma, 61  
Sarig Yughur, 244, 331, 349, 353, 356, 357,  
360, 397, 401, 403, 404, 414,  
420  
Sarikoli, 24, 140, 149, 150, 152, 163, 347,  
397, 404, 419, 442, 445  
Selkup, 33, 34, 151, 355, 363, 364, 369,  
370, 375, 379, 384, 396, 400,  
413, 420, 444, 447  
Serbi-Mongolic, 27  
Sheko, 61, 62, 82  
Shira Yughur, 217, 218, 232, 243, 245, 250,  
331, 400, 419, 443, 446  
Shirongolic, 27, 217–219, 231, 232, 238,  
242, 243  
Shodon, 166, 190  
Shom Peng, 44  
Shor, 332, 349, 350, 353, 356, 360  
Shuri, 15, 58, 106, 166, 175–180, 185,  
187, 188, 190–193, 195, 198, 244,  
401–403, 419, 442, 445  
Sibe, 7, 31, 223, 229, 284–286, 299, 307–  
310, 312–314, 317, 324, 327, 328,  
330, 331, 397, 404, 420, 444,  
447  
Siberian Turkic, 350, 351, 355, 360, 370,  
413  
Sinitic, 16, 25, 26, 28, 29, 33, 41, 43, 46, 47,  
52, 76, 86, 242, 251, 254–256,  
260, 269, 274, 354, 413, 415  
Sino-Tibetan, 10, 28, 41, *see* Trans-  
Himalayan  
Sino-Tibetan-Austronesian, 30  
Sirenikski, 21, 22, 403, *see* Sirenik  
Slavey, 57, 384, 385  
Slavic, 22–24, 35, 64, 140, 143–145, 151,  
152, 158, 165, 411  
Sogdian, 24, 140, 148, 151, 152, 162, 163,  
355, 400, 407, 413  
Solon, 27, 32, 44, 52, 220, 229, 231, 247,  
284, 285, 289, 291–295, 302,  
312–314, 316–318, 320, 322,  
323, 404, 411, 413, 420, 443, 446  
Sonai, 166, 185, 186, 188, 190, 193, 196,  
403, 419, 442, 445  
Sumerian, 10  
Surzhyk, 24  
Tai-Kadai, 30, 41  
Taimyr Pidgin Russian, 16, 140, 147, 148,  
152, 159, 371  
Tajik, 24, 149, 163, 404  
Tangut, 29, 52, 255, 262, 270, 271, 273,  
283

*Language index*

- Tangwang, 16, 38, 52, 255, 267, 273, 278,  
404, 414, 420, 443, 446
- Tarama, 166, 176, 180, 181, 188, 190, 198,  
401
- Tatar, 32, 332, 335, 336, 340, 353, 356, 358,  
360, 420, 444, 447
- Tavgy, *see* Nganasan
- Teiwa, 94
- Tibetan, 16, 28, 29, 48, 52, 100, 255, 268–  
270, 272, 281, 411
- Tibetic, 16, 28–30, 52, 240, 250, 255, 266,  
268–272, 281, 283
- Tibeto-Burman, 28
- Timor-Alor-Pantar, 81, 94, 412
- Tlingit, 384, 385
- Tocharian, 23, 38
- Tocharian A, 23, 140, 150–153, 284, 400
- Tocharian B, 23, 80, 150–152, 165, 406,  
407, 411
- Tocharian C, 23
- Tofa, 334, 347, 353, 355–357, 400, 420,  
444, 447
- Trans-Himalayan, 10, 15, 16, 18, 28–30,  
41, 46, 47, 50, 66–68, 72, 73, 75,  
82, 88, 140, 255, 273, 375, 396
- Trans-New Guinea, 67, 68
- Transeurasian, 44, 50, 51
- Tsezic, 67
- Tshangla, 67, 71
- Tsuken, 39, 166, 175, 176, 180, 188, 190,  
193, 401
- Tumshuqese, *see* Saka
- Tungusic, 7, 13, 16, 18–20, 22, 26, 27,  
30–32, 35, 38, 43–52, 76, 77,  
79, 86, 93, 94, 96, 98, 103, 116,  
117, 119, 119<sup>2</sup>, 120, 124, 128, 130,  
137, 205, 213, 220, 221, 223, 225,  
229–231, 242, 244, 246, 247,  
284–287, 290, 296, 299, 301,  
303, 306–308, 311, 312, 314–  
319, 325–328, 331, 333, 365,  
374, 376, 392, 397, 404–406,  
410, 413, 415–417, 436
- Tunni, 58
- Turkic, v, 11, 16, 18, 27, 31–35, 38, 40,  
43, 44, 46, 48–52, 100, 218,  
219, 222, 230, 235, 238, 242,  
244, 245, 247, 255, 261, 263,  
265, 280, 331–334, 336, 343,  
344, 347, 351–357, 369, 370,  
375, 379, 380, 387, 393, 396,  
398, 400, 401, 404–406, 412–  
415, 437
- Turkish, 11, 12, 40, 46, 331–336, 347, 352–  
354, 356, 358, 360, 400, 405
- Turkmen, 32, 332
- Tuvan, 15, 49, 235, 332, 334, 345–347,  
351–353, 355–357, 360, 400,  
404, 407, 413, 420, 444, 447
- Tuyuhun, 52
- Tzeltal, 101
- Udegheic, 30, 284, 285, 300, 312, 314, 315,  
325, 326, 413, 417
- Udihe, 13, 31, 94, 285, 297–300, 300<sup>26</sup>,  
301, 302, 305, 306, 312–316,  
324–326, 402–404, 413, 416,  
417, 420, 443, 446
- Udmurt, 364
- Uilta, v, 19, 20, 25, 39, 94, 96, 116, 117,  
120, 231, 285, 289, 294, 302–  
306, 312–314, 317, 318, 325, 326,  
398, 404, 413, 420, 444, 447
- Ukrainian, 15, 24, 140, 141, 145, 146, 151,  
152, 159, 397, 400, 404, 408,  
409, 411, 419, 442, 445
- Ulcha, 20, 228, 285, 302, 312–314, 316–  
318, 397, 398, 420, 444, 447
- Ulchi, *see* Ulcha
- Ura, 166, 173, 174, 176, 186, 188, 190, 191,  
193, 397
- Uralic, 7, 10, 13, 18, 33, 35, 40, 44–47, 51,  
55, 63, 64, 81, 88, 331, 355, 363–  
366, 379, 380, 405, 437
- Urarina, 58
- Uyghur, 15, 16, 24, 27, 31–33, 49, 140,  
149, 218, 263, 265, 267, 331, 332,

- 337, 340–344, 347, 349, 353,  
356, 360, 396, 397, 401, 403–  
405, 413–415, 420, 444, 447
- Uygur-Karluk, 32, 33, 331, 332
- Uzbek, 15, 32, 94, 149, 332, 340, 344, 345,  
353, 354, 356, 359, 360, 404
- Veps, 364
- Votic, 364
- Wadul, *see* Yukaghir
- Wakhi, 24, 149, 163
- Wari', 59, 77
- West Greenlandic, 45, 100, 128
- Wutun, v, 16, 38, 52, 86, 100, 168, 242, 255,  
265–267, 273, 278, 280, 401,  
403, 404, 410, 414, 415, 420,  
443, 446
- Yaghnobi, 24, 162, 163
- Yakut, 15, 31–33, 35, 49, 244, 332, 334,  
335, 352, 353, 355–357, 360,  
387, 397, 400, 401, 404, 420,  
444, 447
- Yéli Dnye, 61, 62, 101
- Yenisei-Ostyakic, 377
- Yenisei-Samoyed, *see* Enets
- Yeniseic, 15, 16, 18, 31, 33–35, 45, 47, 377,  
381–383, 385, 435
- Yiddish, 24, 122, 140–145, 151–154, 156,  
157, 369, 396, 397, 404, 419,  
442, 445
- Yilan Creole, 16, 25, 165, 166, 170, 187–  
190, 193, 198–200, 419, 442,  
445
- Yugh, 34, 377, 379–382
- Yukaghir, 33, 35, 290, 296, 351, 375, 386–  
390, 390<sup>38</sup>, 391–394, 397, 398,  
403, 404, 407–409, 420, 444,  
447
- Yukaghiric, 15, 18, 21, 31–35, 45–48, 50,  
51, 76, 124, 132, 136, 189, 296,  
352, 354, 355, 386, 389, 392,  
393, 396, 405, 406
- Yukcin, 25, 200, 201, 210
- Yupik, 12, 21, 22, 76, 94, 124, 128–132,  
134–137, 374, 395, 396, 408,  
409, 419
- Yurak, *see* Nenets
- Yurats, 364
- Yuwan, 58, 166, 174–178, 180, 181, 183, 187,  
188, 190, 193, 195, 196, 198, 401,  
419, 442, 445
- Zhongu, 29, 255, 255<sup>22</sup>, 270, 272, 273, 281,  
420, 443, 446
- †Ākhoe Hailom, 67



# Subject index

- A-not-A question, 60, 65, 258, 259, 261  
Abdal, 331  
accretion zone, *see* residual zone  
acquisition, 3, 7, 43, 79, 87, 415, 416  
action, 3, 4, 90, 91, 99, 266, 335  
admixture, 19, 31  
affordances, 4  
affords, 99  
agriculture, 436  
Alaska, 8, 18, 21, 22, 24, 34  
Altai, 8, 11, 23, 34, 48, 153, 156, 331, 353,  
355–357, 360, 418, 435, 444,  
447  
Altaic, 45, 50, 51  
Altaicization, 43  
alternative question, 5, 56, 57, 59, 60,  
65–68, 70, 71, 73, 74, 91–93,  
95, 105, 107, 108, 116, 130, 131,  
143, 144, 147–151, 153, 172, 183,  
184, 205, 206, 211, 212, 221, 222,  
224, 229–234, 236, 237, 240,  
258–261, 263, 270, 287, 289,  
290, 292–295, 297, 298, 301,  
306, 307, 309, 333, 336, 337,  
340, 345, 349–351, 354, 365–  
369, 378, 379, 386, 387, 396–  
398, 412, 414, 415  
alternative questions, 55  
Amdo, 8, 11, 33, 48, 268, 270, 281, 282,  
396, 397, 415, 418, 435  
Amdo Sprachbund, 16, 27, 29, 40, 52, 255,  
269, 281, 360, 414, 415, 436  
Amur, 8, 11, 18, 20, 30, 31, 48, 52, 114–117,  
119, 402, 411, 413, 419, 435, 442,  
445  
analysis, 12, 13, 61, 77, 89, 116, 128, 135,  
137, 153, 189, 191, 201, 206–208,  
216, 221, 225, 229, 231, 237, 238,  
241, 242, 244, 247, 254, 255,  
262, 265, 274, 275, 290, 299,  
304, 307, 347, 351, 369, 371, 377,  
378<sup>36</sup>, 382, 387, 389  
analyzability, 77, 78, 87, 283, 316  
anatomically modern humans, 10  
Ancient North Eurasians, 437  
answer, 58–60, 68, 92, 93, 96, 97, 99, 101,  
116, 124, 130, 197, 267, 366  
Anthropocene, 2  
anticipation, 91, 94, *see* predictions  
anticipation rule, 100, 267  
Araxes-Iran Linguistic Area, 40  
Arctic Ocean, 8, 31  
areal linguistics, 4, 14, 39, 47, 435  
areal typology, 1  
Asia, 8, 11, 12, 17, 21, 22, 26, 32, 34, 42, 45,  
50, 201, 396  
atrophied, 87  
Australasians, 17  
Australasia, 437  
Baikal, 8, 10, 18, 31, 35, 47, 296, 437  
basic question words, 77, 79, 275  
basic semantic categories, 82, 89  
Bering Strait, 8, 11, 21, 22  
Beringia, 17, 34, 437  
bilingualism, 6, 48, 52  
borrowing, 4, 5, 38, 39, 51, 213, 220, 311,  
354, 355, 412, 414, 417  
brain, 90, 91, 438  
case, 42, 45, 52, 78, 83, 84, 88, 111, 119, 120,  
127, 147, 156, 157, 159, 163, 169,

## Subject index

- 193, 206, 208, 215, 247, 250,  
254, 281, 316, 317, 320, 322–  
324, 329, 374, 382, 389–391,  
393, 408
- causal frames, 5, 438
- China, 4, 7, 8, 10, 11, 13, 18, 24, 26, 27,  
29, 30, 32, 34, 41, 43, 47, 49, 52,  
55, 68, 160, 163, 199–202, 218–  
220, 255, 257–260, 262, 263,  
280, 286, 299, 307, 331, 335–  
340, 346, 377, 396, 398, 402,  
418, 435, 437
- Chukotka, 8, 11, 17, 18, 20, 24, 396
- climate, 2–4, 8, 436, 438
- cognition, 5, 438
- cognitive ecology, 4
- Cognitive Linguistics, 437, 438
- Cognitive Science, 1, 98
- cognitive typology, 4
- collative, 5, 95, 96, 439
- combination, 48, 72, 75, 77, 79, 93, 95,  
132, 143, 148, 172, 175, 177, 202,  
214, 215, 229, 242, 259, 261,  
263, 275, 278, 280, 326, 328,  
334<sup>33</sup>, 340, 341, 344, 365, 382,  
391, 399, 412
- communicative motives, 92
- complexification, 416
- complexity, 5, 12, 13, 75, 78, 87, 96
- conception, 3, 52, 91
- conceptual space, v, 70, 71, 82–84, 86,  
398, 407
- confirmation, 59
- conjunct/disjunct, *see* evidentiality
- construal, 92
- contact, 6, 7, 19–23, 25, 26, 31–35, 37–39,  
46, 49, 50, 87, 124, 169, 191, 220,  
247, 328, 375, 415–417, 438
- content question, 5, 6, 12, 39, 54, 56–60,  
65, 67, 68, 70, 71, 93, 104, 105,  
107–109, 116, 121–124, 129, 130,  
132, 136, 143, 144, 146, 149, 150,  
167, 168, 170, 172, 175–187, 204,  
205, 211, 219, 222, 223, 227–  
229, 231, 232, 242, 257, 260,  
263, 265, 267, 270, 273, 291,  
293, 294, 296, 298, 300, 302,  
303, 305, 307, 309, 334, 335,  
339, 345, 347, 348, 350–352,  
354, 365–369, 379, 380, 384–  
387, 390, 396, 398, 399, 401,  
412, 414, 416
- continuum, 7, 128, 165, 200
- conventionalization, 3
- converb, 60, 83, 84, 118, 176, 198, 246, 315,  
325, 326, 328
- convergence, 86
- convergent evidence, 14, 22
- coordination, 65, 95, 436
- creoles, 38, 55, 87, 416
- curiosity, 5, 93, 96–99, 101, 170, 439
- declarative sentence, 56, 140, 144, 146,  
206, 230, 257, 387, 399
- demonstratives, 6, 53, 59, 76, 77, 80, 84,  
86, 88, 89, 119, 127, 164, 165, 187,  
189, 193, 195, 215, 244, 247, 275,  
312, 315, 317, 319, 323–325, 329,  
334, 355, 374, 394
- Denisovans, 10
- diachronic, 1–3, 6, 89, 90, 128, 175, 256,  
284, 410
- dialogical array, 99, 101, 439
- disjunction, 57, 65–68, 72, 73, 116, 130,  
131, 143, 146–149, 151, 183, 187,  
205, 260, 261, 290, 292, 298,  
309, 310, 310<sup>29</sup>, 312, 333, 336,  
337, 342, 349, 350, 352, 354,  
378, 387, 388, 396–398, 425,  
441
- double marking, 66, 104, 108, 151, 183,  
184, 220, 263, 290, 298, 354,  
366, 378, 396, 398, 425, 441,  
444<sup>1</sup>
- dual, 127, 128, 135
- East Asia, 10, 19, 31

- echo questions, 172  
 ecolinguistics, 1  
 ecological commitment, 437  
 Ecological Psychology, 90, 93, 98  
 ecological theory of questions, 90, 439  
 ecological typology, 14, 437  
 ecology, 2, 4–6, 12, 90, 438  
 ellipsis, 68  
 embodied simulation, 59, 91, 97, 98, 100  
 emphasis, 11, 53, 95, 122, 123, 221, 257, 395  
 enchronic, 2–4, 37, 53, 90, 99  
 entrenched situated conceptualization, 92, 93  
 entrenchment, 3  
 ergative, 85, 128, 389  
 Eurasia, 41, 42  
 Europe, 8, 24, 41, 42, 44, 48, 50, 51, 61, 63, 217, 263, 397, 437  
 evidentiality, 75, 99, 267, 335  
 evolution, 3, 4  
 exploration, 94, 99, 439, *see* exploratory behavior  
 exploratory behavior, 5, 93, 96, 99, 101, 439  
 eye contact, 99  
 Far Eastern Federal district, 8  
 focus, 54, 56–61, 64, 68, 70–72, 74, 76, 92, 93, 107, 114, 115, 128, 130, 144, 146, 153, 168, 169, 174–176, 178–187, 211, 225, 227, 231, 238, 258, 259, 269, 286, 289, 296, 297, 301, 302, 304, 333, 340, 348, 351, 354, 368, 379, 388–392, 397–399, 401, 412, 436  
 focus question, 56–60, 68, 74, 75, 93, 94, 115, 123, 128, 130, 141, 142, 145, 167, 168, 170, 171, 174, 175, 178, 182, 183, 185, 187, 206, 221, 223, 224, 227, 228, 231, 238, 258, 259, 262, 266, 287, 291, 297, 301, 302, 304, 333, 336, 347, 351, 370, 387, 397, 399  
 front rounded vowels, 45–47, 406, 437  
 fronting, 56, 64, 167, 312  
 functional domain, 72, 398, 436  
 fusion, 72, 399  
 Gansu, 52, 233, 255, 263<sup>23</sup>, 414  
 gazing behavior, 53, 101  
 gender, 75, 156, 159, 165, 382, 403, 411  
 genetic bias, 1  
 genetics, *see* human genome  
 gestalt, 97  
 Gestalt Psychology, 97  
 gestures, 90, 99  
 glossing, 12, 141<sup>5</sup>, 167, 207, 299  
 Gobi, 10  
 grammar of questions, 4, 5, 7, 12, 13, 16, 38, 54, 75, 103, 126, 331, 395, 412, 414, 417, 436  
 grammaticalization, 55, 59, 72, 73, 83, 86, 169, 221, 229, 247, 259, 399, 401  
 Greater Himalayan Region, 40  
 Greater Manchuria, 11  
 Guiyang, 278  
 Gulf of Bohai, 8, 11  
 head shake, 53  
 Heilongjiang, 41  
 hierarchy, 55, 78, 79, 93, 101  
 hierarchy of specificity, 93  
 Himalayas, 18, 28, 48  
 Hokkaidō, 8, 18, 19, 48, 103, 104, 107, 108  
 Holocene, 17, 35  
 homeland, 18, 20, 21, 27, 28, 30, 33, 34, 41  
 Homo erectus, 10  
 Honshū, 18, 19, 103  
 huh, 37, 55  
 human genome, 7, 22, 43  
 hunter-gatherer, 5, 25, 55  
 imaginative capacity, 92  
 imperative, 123, 124  
 in situ, 147, 312  
 Indigirka, 8, 18  
 indirect question, 56

*Subject index*

- indirect questions, 118, 144, 366  
Indoeuropean, 55  
inflection, 83, 86, 88, 89, 157, 174, 175, 410  
informing, 92, 93  
interaction, 3, 33, 34, 38, 52, 56, 61, 62,  
65, 66, 68, 72, 75, 77, 90, 99, 101,  
123, 175, 182, 238, 241, 266, 267,  
269, 397, 399, 415  
interrogation, 53, 56, 106  
interrogative, 5–7, 13, 38, 42, 47, 53, 55,  
56, 58–60, 64, 66, 73, 76–84,  
86, 87, 89, 90, 100, 104, 106,  
108, 109, 111, 115, 117, 119, 120,  
122, 124, 126, 128–132, 134–  
137, 143, 144, 146–148, 150, 151,  
153, 156, 157, 159, 162, 164,  
165, 169, 171, 177, 179, 183,  
184, 187, 189, 191–193, 195–199,  
201–203, 205, 207–209, 211,  
212, 214–216, 219, 221, 226, 227,  
230, 232, 233, 236, 238, 240,  
242, 244, 246, 247, 250, 252–  
254, 257, 259, 264, 265, 269,  
275, 278, 280, 281, 283, 294–  
301, 303<sup>27</sup>, 304<sup>28</sup>, 305–307, 311,  
312, 314–320, 322, 324–328,  
337, 354, 355, 357, 358, 360,  
365, 366, 369, 371, 373–375,  
378–380, 382–384, 389, 390,  
392–394, 398, 405, 406, 414–  
417, 436, 447<sup>4</sup>  
interrogative systems, 7, 87, 197, 278,  
320, 366, 371, 375, 381, 382, 407,  
410, 416  
interrogative verb, 60, 119<sup>2</sup>, 127, 177, 198,  
214, 238, 246, 247, 250, 251,  
298, 315, 348, 355, 358, 365, 391  
interrogativity, 53, 63, 99, 100  
intonation, 4, 37, 55, 56, 62, 63, 65, 66, 68,  
71<sup>1</sup>, 76, 104, 106, 115, 121, 122,  
130, 131, 140, 141, 143, 144, 146–  
148, 150, 167, 170, 172–174, 179,  
184, 185, 210, 223, 227, 257, 268,  
269, 278, 286, 293, 295, 296,  
300–302, 333, 336, 338, 339,  
341, 346, 348, 351, 365, 366,  
368, 369, 380, 386, 388, 395,  
399, 401, 436, 438, 441  
isolate, *see* language isolate  
Jōmon, 19, 25  
Japan, 8, 10, 11, 19, 24, 25, 46, 49, 51, 201,  
259, 262, 396, 418, 435  
Jurchen, 49, 284–286, 328  
juxtaposition, 61, 65, 66, 130, 184, 259,  
290, 294, 379  
K-interrogatives, 6, 41, 316, 405, 406, 432  
kakari musubi, 168, 174, 183, 187, 189, 391,  
399  
Kamchatka, 11, 19, 20, 24, 31, 48, 49, 52,  
396, 418  
Kamchatkan isthmus, 18  
Karakorum, 8  
KIN-interrogative, 6, 47, 111, 128, 153,  
246, 371, 405, 431  
Kolyma, 8, 35, 444, 447  
Korea, 8, 10, 11, 18, 25, 26, 30, 31, 51, 201,  
396, 418, 435  
Korean Peninsula, 10, 25, 27, 200, 201  
Krorän, 23  
Kucha, 23  
Kunlun, 8  
Kuril, 19, 20, 48, 103, 104, 109  
landscape roughness, 436  
language change, 3, 38, 414  
language contact, 4, 5, 7, 14, 31, 37, 38,  
47, 52, 64, 71, 319, 331, 395, 405,  
414–417, 438  
language death, 38  
language diversity, 1, 43, 45, 47, 435, 436  
language family, 15, 16, 20, 22, 25, 26, 28,  
30, 31, 33, 34, 38, 47, 50, 51, 103,  
113, 120, 128, 165, 217, 377, 398,  
416  
language isolate, 15

- language shift, 4, 5, 31  
 language spread, 4, 30, 45, 47, 48  
 laryngeal sounds, 37, 38  
 Lena, 8, 18, 35, 47, 296, 437  
 Liao, 8, 27, 30  
 Liao-dynasty, 27  
 Liaoning, 27  
 lingua franca, 19, 24  
 linguistic area, 27, 39, 43, 49–52, 414  
 linguistic diversity, 1–4, 40, 44, 45, 167, 435–438  
 linguistic typology, 1, 3, 13  
 looking, 90, 99, 101, 328  
 Luoravetlan, *see* Chukotko-Kamchatkan  
  
 m-T-pronouns, 6, 42, 43, 405, 406  
 Mainland Southeast Asia, 19, 40, 41, 43, 44, 50, 435–437  
 Manchuria, 8, 11, 20, 24, 25, 27, 30–32, 47, 49–52, 217, 284, 331, 402, 418  
 Manchurian Plain, 10  
 markedness, 84, 395  
 matter of degree, 15, 40, 77  
 mental access, 95  
 microgenetic, 2–5, 90, 99  
 migration, 17, 21, 22, 25, 34, 48, 331  
 mixed languages, 38, 416  
 Mongolia, 8, 10, 11, 16, 18, 27, 32, 34, 47, 49–51, 159, 217, 218, 225, 331, 377, 418, 435  
 Mongolian Plateau, 10  
 morphology, 42, 43, 49, 111, 124, 128, 137, 177, 247, 275, 319, 365  
  
 nasals, 49, 50, 227, 291  
 natural boundary, 8  
 natural ecology, 4  
 natural selection, 96  
 NEA, 1, 5–8, 10–18, 21, 23, 24, 29, 31–35, 38–50, 52, 53, 56, 58, 60–63, 65, 69, 70, 73–76, 79–81, 83, 86, 88, 98, 103, 121, 123, 124, 128, 136, 137, 140, 151, 189, 255, 256, 271–273, 275, 278, 290, 331, 344, 354, 355, 368, 375, 377, 381, 385, 389, 395–408, 412, 413, 415–418, 435–438, *see* North East Asia  
 Neanderthals, 10  
 negation, 59, 61, 66, 72, 73, 86, 259, 262, 264, 287, 353, 365, 387, 401, 436  
 negative alternative question, 59, 61, 73, 225, 257, 259, 260, 265, 287, 290, 294, 306, 309, 337, 342, 343, 368, 369, 379  
 negative polar question, 58, 177, 378  
 niche construction, 2, 96  
 nominalization, 72, 105, 169  
 nominalizations, 107  
 North Korea, 8, 49  
 Northeast Asia, 1, 4, 7, 8, 10–14, 23, 26, 30, 37, 40, 43, 45, 93, 126, 128, 183, 201, 218, 224, 273, 334, 335, 340, 354, 364, 369, 370, 373, 381, 382, 399, 418, 435–438  
 Northeast Eurasia, 8  
  
 Okhotsk culture, 20  
 Okhotsk people, 19  
 ontogenetic, 2–4, 93  
 opaque, 69, 70, 87, 109, 128, 162, 275, 280, 382, 408  
 open alternative question, 59, 83, 260, 295  
 Ordos Plateau, 10  
 organism-environment system, 2, 4, 90, 97, 99, 438, 439  
  
 Pacific Ocean, 10  
 Pacific Rim, 8, 46, 435  
 Paleo-Eskimos, 17  
 Pamir, 8, 24, 48  
 Pamir-Hindukush Sprachbund, 40  
 pastoralism, 436  
 perception, 3, 90, 91, 95

## Subject index

- perception-action cycle, 93  
phylogenetic, 2–4, 44, 48, 93, 140, 436  
phylogenetic diversity, 1, 43–45, 47, 435  
pidgins, 16, 38, 55, 147  
plural, 88, 127, 134, 135, 137, 156, 157,  
196, 252, 278, 280, 281, 296, 341,  
360, 373, 382, 391, 392  
pointing, 31, 89, 90  
polar question, 4, 7, 37, 38, 54–58, 61,  
63–65, 67, 68, 70, 72–74, 76,  
104, 108, 109, 114, 121, 128, 130,  
136, 140–142, 144–147, 149, 150,  
168, 171, 172, 174, 176, 177, 179,  
182–184, 187, 204, 205, 212, 219,  
223, 225, 227, 228, 230, 237,  
240, 242, 256, 257, 263, 265–  
267, 270, 273, 286–288, 295,  
297, 299, 300, 302, 309, 336,  
350–352, 354, 364, 366, 368,  
369, 380, 387, 388, 390, 395–  
399, 436, 438  
polar questions, 115, 116  
Pontic-Caspian steppe, 22, 437  
population density, 436  
possession, 319, 340  
precipitation, 436, 438  
predictions, 91, 92, 98–100, 439  
prehistory, 7, 12, 17, 19  
proto-languages, 6, 10, 28, 46, 47, 50, 221,  
375, 405  
prototypical questions, 92  
Punuk, 22  
  
Qarashähär, 23  
Qianlong, 331  
Qinghai, 32, 47, 52, 278  
Qinling, 8  
question marker, 4, 7, 38, 39, 59, 64, 66,  
68, 70–74, 76, 100, 106–108, 115,  
117, 122, 123, 129, 130, 132, 141,  
143, 146–151, 158, 159, 167, 169–  
173, 175–183, 185–187, 206, 208,  
211, 212, 214, 219–227, 229–  
234, 236, 238, 240, 242, 256–  
263, 265, 270–272, 274, 287,  
289, 290, 292, 294, 295, 298,  
300, 301, 304, 305, 307–310,  
333–337, 339–354, 357, 364–  
370, 378–380, 384, 387, 388,  
392, 396–398, 401, 402, 414,  
415  
question marking, v, 4, 6, 7, 13, 54–58, 60,  
61, 63, 65, 66, 68, 72, 74–76, 82,  
103, 104, 109, 123, 132, 136, 141,  
143, 149, 167, 173, 183, 189, 201,  
204, 212, 224, 237, 238, 240,  
242, 256, 260, 271, 289, 290,  
294, 333, 336, 340, 352, 366,  
367, 370, 384, 385, 389, 395–  
398, 401, 402, 415, 436, 438  
question tag, 61, 69, 73, 95, 116, 146, 150,  
151, 230, 258, 259, 261, 305, 343,  
347, 380<sup>37</sup>, 397  
question type, 59–61, 95, 204, 290, 302  
question-response sequences, 3, 53, 172  
questions, v, 3, 5, 13, 14, 24, 37–39, 53,  
54, 56–61, 63–74, 76, 90–95, 97,  
99–101, 104–107, 116, 121, 124,  
128, 130–132, 143, 144, 146, 148,  
151, 167, 169–174, 176–181, 183–  
185, 187, 197, 198, 202, 206, 207,  
209, 211–213, 219, 223, 225–  
227, 230–232, 234, 235, 237–  
240, 255–263, 265–270, 273,  
275, 280, 287, 290–302, 304,  
306–309, 312, 333, 334, 336,  
337, 341–350, 352–354, 364–  
370, 379, 380, 386, 388–390,  
395–399, 412, 415, 436, 438,  
439  
Qäwrighul, 23  
  
reconstruction, 28, 46, 51, 125, 191, 274,  
312, 314, 316, 318, 367, 393  
reduction, 7, 43, 87, 270, 416  
reduplication, 88, 281, 306, 360  
reference point, 95, 130, 333, 334  
reindeer, 2, 20, 21, 48, 438

- reinforcement, 86, 410  
 replacement, 86  
 request, 56, 57, 185, 439  
 requesting, 92, 93  
 residual zone, 40  
 resonance, 6, 77, 85, 86, 111, 117, 125, 136,  
     137, 153, 164, 191, 197, 214, 215,  
     247, 250, 275, 283, 316, 325,  
     355, 358, 360, 371, 375, 382,  
     393, 394, 406, 410  
 response, 97, 124, 167, 412  
 rhetorical, *see* rhetorical question  
 rhetorical question, 55, 118  
 Rhetorical questions, 116, 148  
 rhetorical questions, 233, 369  
 rice, 43  
 riddle, 54, 96  
 river density, 436  
 Russia, 8, 10, 11, 24, 27, 31, 32, 49, 51, 218,  
     336, 347  
 Ryūkyūan Islands, 25, 39, 48, 167, 396,  
     418, 435  
 Sakhalin, 11, 19, 20, 26, 31, 39, 48, 103, 104,  
     108, 109, 111, 114, 115, 117, 120,  
     201, 288, 320, 402, 413, 418, 419,  
     442, 445  
 schematicity, 93  
 Sea of Okhotsk, 8  
 selection, 59  
 semantic map, 71  
 Semantic Map Connectivity Hypothe-  
     sis, 71  
 semantic scope, 70, 71, 80, 82, 107–109,  
     137, 170, 173, 247, 288, 297, 306,  
     364, 367, 368, 370, 380, 395,  
     398, 406–408, 415  
 sharing, 92  
 Siberia, v, 7, 10, 11, 17, 18, 24, 30–33, 35,  
     48–51, 98, 130, 159, 331, 396,  
     418, 435  
 Siberian Federal district, 8  
 Sichuan, 52  
 Silla, 26  
 similarity, 6, 7, 37–39, 47, 53, 78, 94, 108,  
     191, 221, 247, 262, 360, 388, 392,  
     405, 438  
 simplification, 7, 87, 328, 414–416  
 single marking, 66, 397, 415, 425, 441  
 singular, 75, 78, 100, 119<sup>2</sup>, 125, 129, 132,  
     134, 135, 137, 156, 254, 296, 300,  
     301, 337, 352, 378, 392  
 Sinocentric view, 10  
 smallpox, 35  
 social ecology, 5  
 sociocultural ecology, 4  
 South Korea, 24, 49, 206  
 Southeast Asia, 8, 19, 28, 43, 396, 437  
 specificity, 60, 93, 94  
 speech act, 3, 50, 172, 202  
 split type, 75, 100, 399  
 sprachbund, 39, *see* linguistic area  
 spread zone, 40, 418  
 steppes, 10  
 structural diversity, 1, 4, 13, 37, 47, 48,  
     436, 437  
 submorpheme, 6, 77, 274, 371  
 subsistence, 21, 436  
 suffixing, 42, 49, 50  
 Sunggari, 30  
 superstrate, 49  
 symbolic ecology, 4, 438  
 synchronic, 1–3, 69, 70, 90, 109, 128, 163,  
     316, 323, 374, 418, 435  
 tag question, 60, 61, 69, 70, 94, 106, 116,  
     131, 143, 144, 149, 172, 175, 176,  
     205, 206, 226, 227, 229, 230,  
     235, 237, 257, 259, 264, 269,  
     290, 292, 298, 304, 308, 310,  
     333, 334, 339, 344, 346, 370,  
     387, 388  
 tag questions, 55  
 Taiga, 8, 10  
 Taiwan, 8, 13, 16, 25, 55, 68, 165, 185, 198,  
     199, 199<sup>18</sup>  
 Taklamakan, 10  
 Tarim, 23, 140

*Subject index*

- Tatary, 10  
temperature, 436  
tendency, 7, 37, 38, 49, 79, 87, 94, 101, 151,  
273, 338, 382, 384, 397, 408,  
412  
thinking, 99  
Thule, 17, 22  
Tianshan, 8  
Tibetan Plateau, 8, 48  
time scales, 3, 90, 436–438  
tones, 64, 438  
Topic question, 340, 344  
topic question, 131, 168, 181, 258  
Topic questions, 147  
toponyms, 19, 34  
transparency, 7, 87, 416  
transparent, 67, 70, 77, 87, 220, 221, 274,  
275, 328, 382, 417  
tropics, 4, 438  
tundra, 8, 11  
Turfan, 23  
turn-taking, 3, 55  
  
ultra-social, 99  
uncertainty, 92, 93, 95, 96, 170, 354  
universal, 3, 13, 37, 53, 56, 63, 69–72, 74,  
79, 94, 99, 397–399, 435  
universal questions, 37  
Ural, 11, 40  
Ural-Altai, 47, 51  
urheimat, *see* homeland  
Ussuri, 30, 31, 302  
  
velar nasal, 39, 49, 50, 304, 312, 320  
Volga-Kama, 40  
vowel harmony, 39, 47, 49, 203, 221, 292,  
304, 307, 338, 346, 351  
  
wagon, 22  
Western Siberian Lowland, 8, 40  
wh-movement, 56  
What is your name?, 37, 81, 82, 154, 181,  
199, 257, 267, 294, 347  
wheat, 43  
  
wheel, 22, 438  
word class, 76, 80, 84, 88  
word order, 7, 41, 43, 49, 52, 63, 64, 66,  
70, 140, 141, 143, 144, 144<sup>8</sup>, 167,  
290, 348, 351, 364, 414, 415, 422  
World Atlas of Language Structures, 14,  
395  
  
Xiaohe, 23  
Xinjiang, 8, 11, 26, 31, 41, 47, 223, 278,  
294, 331, 339, 340, 344, 396,  
397, 402  
Xiongnu, 34, 47, 377  
Xixia, 29  
  
Yalu, 8  
Yamnaya, 437  
Yangtze, 41, 43  
Yellow River, 8, 11, 28, 47, 435  
Yenisei, 8, 11, 24, 31, 33–35, 40, 332, 396,  
418, 435, 437  
yukar, 103  
Yunnan, 41, 48, 255, 272



# Did you like this book?

This book was brought to you for free

Please help us in providing free access to linguistic research worldwide. Visit <http://www.langsci-press.org/donate> to provide financial support or register as a community proofreader or typesetter at <http://www.langsci-press.org/register>.





# A typology of questions in Northeast Asia and beyond

This study investigates the distribution of linguistic and specifically structural diversity in Northeast Asia (NEA), defined as the region north of the Yellow River and east of the Yenisei. In particular, it analyzes what is called the grammar of questions (GQ), i.e., those aspects of any given language that are specialized for asking questions or regularly combine with these. The bulk of the study is a bottom-up description and comparison of GQs in the languages of NEA. The addition of the phrase *and beyond* to the title of this study serves two purposes. First, languages such as Turkish and Chuvash are included, although they are spoken outside of NEA, since they have ties to (or even originated in) the region. Second, despite its focus on one area, the typology is intended to be applicable to other languages as well. Therefore, it makes extensive use of data from languages outside of NEA. The restriction to one category is necessary for reasons of space and clarity, and the process of zooming in on one region allows a higher resolution and historical accuracy than is usually the case in linguistic typology. The discussion mentions over 450 languages and dialects from NEA and beyond and gives about 900 glossed examples. The aim is to achieve both a cross-linguistically plausible typology and a maximal resolution of the linguistic diversity of Northeast Asia.

ISBN 978-3-96110-102-3



9 783961 101023