

Languages of the World

Third Edition

Are you curious to know what all human languages have in common and in what ways they differ? Do you want to find out how language can be used to trace different peoples and their past? Then this book is for you! Now in its third edition, it guides beginners through the rich diversity of the world's languages. It presupposes no background in linguistics, and introduces the reader to linguistic concepts with the help of problem sets, end-of-chapter exercises, and an extensive bibliography. Charts of language families provide geographical and genealogical information, and engaging sidebars with demographic, social, historical, and geographical facts help to contextualize and bring languages to life. This edition includes a fully updated glossary of all linguistic terms used, new problem sets, and a new section on cartography. Supplementary online materials include links to all websites mentioned and answers to the exercises for instructors.

ASYA PERELTSVAIG is a native speaker of Russian and a fluent speaker of three other languages. For the last fifteen years, she has taught courses in linguistics at Yale, Cornell, Stanford, and several other universities around the world. Her academic interests include languages, history, genetics, and the relationship between the three.

CAMBRIDGE

Cambridge University Press & Assessment
978-1-108-47932-5 — Languages of the World
Asya Pereltsvaig
Frontmatter
[More Information](#)

LANGUAGES OF THE WORLD

An Introduction
Third Edition

ASYA PERELTSVAIG
Independent scholar



CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE
UNIVERSITY PRESS

University Printing House, Cambridge CB2 8BS, United Kingdom

One Liberty Plaza, 20th Floor, New York, NY 10006, USA

477 Williamstown Road, Port Melbourne, VIC 3207, Australia

314–321, 3rd Floor, Plot 3, Splendor Forum, Jasola District Centre, New Delhi – 110025, India

79 Anson Road, #06–04/06, Singapore 079906

Cambridge University Press is part of the University of Cambridge.

It furthers the University's mission by disseminating knowledge in the pursuit of education, learning and research at the highest international levels of excellence.

www.cambridge.org

Information on this title: www.cambridge.org/9781108479325

DOI: 10.1017/9781108783071

© Asya Pereltsvaig 2012, 2017, 2021

This publication is in copyright. Subject to statutory exception and to the provisions of relevant collective licensing agreements, no reproduction of any part may take place without the written permission of Cambridge University Press.

First published 2012

Second edition published 2017

Third edition published 2021

Printed in the United Kingdom by TJ International Ltd, Padstow Cornwall, 2021

A catalogue record for this publication is available from the British Library.

ISBN 978-1-108-47932-5 Hardback

ISBN 978-1-108-74812-4 Paperback

Additional resources for this publication at www.cambridge.org/pereltsvaig3

Cambridge University Press has no responsibility for the persistence or accuracy of URLs for external or third-party internet websites referred to in this publication and does not guarantee that any content on such websites is, or will remain, accurate or appropriate.

Cambridge University Press & Assessment
978-1-108-47932-5 — Languages of the World
Asya Pereltsvaig
Frontmatter
[More Information](#)

To my parents, Freyda and Michael Pereltsvayg, ל”ת

Contents

<i>List of Figures</i>	<i>page x</i>
<i>List of Tables</i>	<i>xi</i>
<i>List of Textboxes</i>	<i>xiii</i>
<i>List of “Focus on” Topics</i>	<i>xiv</i>
<i>Preface to the Third Edition</i>	<i>xv</i>
<i>Acknowledgments</i>	<i>xvii</i>
<i>Abbreviations Used in the Glosses</i>	<i>xviii</i>
1 Introduction	1
1.1 Languages, Dialects, and Accents	5
1.2 Language Families	10
1.3 How Language Families Are Established: Comparative Reconstruction	14
1.4 Linguistic Diversity	22
1.5 How Do Languages Diversify?	26
1.6 Field Linguistics	28
1.7 Focus on: Language Maps	32
Do It Yourself	37
2 Languages of Europe	43
2.1 The Indo-European Language Family	45
2.2 The Indo-European Controversy	55
2.3 Non-Indo-European Languages of Europe	60
2.4 Focus on: Endangered Languages of Europe	65
Do It Yourself	71
3 Languages of Iran and South Asia	75
3.1 Indo-European Languages of Iran and South Asia	76
3.1.1 Iranian Languages	78
3.1.2 Indo-Aryan Languages	81
3.2 Dravidian Languages	87
3.3 Other Languages of South Asia	91
3.4 Focus on: Universals and the Parametric Theory of Language	95
Do It Yourself	103
4 Languages of Northern Eurasia	108
4.1 Finno-Ugric Languages	110
4.2 Other Uralic Languages	119
4.3 Turkic Languages	122
4.4 Other Languages of Siberia	128
4.5 Focus on: Evidential Markers	136
Do It Yourself	140

5 Languages of the Caucasus	146
5.1 Northwest Caucasian Languages	152
5.2 Northeast Caucasian Languages	157
5.3 Kartvelian Languages	163
5.4 Indo-European Languages in the Caucasus	168
5.5 Focus on: Head-Marking vs. Dependent-Marking	172
Do It Yourself	176
6 Languages of the Greater Middle East	181
6.1 Afroasiatic Languages	183
6.2 Semitic Languages	190
6.3 Berber Languages	203
6.4 Focus on: Language Contact	206
Do It Yourself	213
7 Languages of Sub-Saharan Africa	216
7.1 Nilo-Saharan Languages	219
7.2 Niger-Congo Languages	224
7.3 Khoisan Languages	234
7.4 Focus on: Official Languages, Trade Languages, and Creole Languages in Sub-Saharan Africa	240
Do It Yourself	245
8 Languages of Eastern Asia	249
8.1 Sino-Tibetan Languages	251
8.2 Austro-Asiatic Languages	261
8.3 Tai-Kadai Languages	265
8.4 Japanese and Korean	270
8.5 Focus on: Isolating Morphology and Language Change	275
Do It Yourself	280
9 Languages of the South Sea Islands	287
9.1 Discovery of the Austronesian Family and the Austronesian Homeland	292
9.2 Today's Austronesian Languages and the Internal Classification of the Family	295
9.3 Linguistic Properties of Austronesian Languages	301
9.4 Focus on: The Mystery of Malagasy	311
Do It Yourself	319

10 Aboriginal Languages of New Guinea and Australia	324
10.1 Languages of New Guinea	328
10.1.1 Overview of Papuan Languages	329
10.1.2 Areal Papuan Features	331
10.1.3 Tok Pisin	336
10.2 Languages of Australia	339
10.3 Focus on: Is There Such a Thing as a Primitive Language?	347
Do It Yourself	355
11 Native Languages of the Americas	359
11.1 Languages of North America	362
11.2 Languages of Meso-America	381
11.3 Languages of South America	386
11.4 Focus on: The Pirahã Controversy	394
Do It Yourself	400
12 Remaining Issues: Macro Families, Sign Languages, and Constructed Languages	406
12.1 Macro Families	406
12.1.1 The Dene-Yeniseian Hypothesis	408
12.1.2 Altaic and Ural-Altaic Macro Families	409
12.1.3 The Nostratic and Eurasiatic Hypotheses	414
12.1.4 Other Hypothesized Macro Families	421
12.2 Sign Languages	425
12.3 Constructed Languages	434
Do It Yourself	441
<i>Glossary</i>	443
<i>Bibliography</i>	449
<i>Index of Languages</i>	471
<i>Index of Terms</i>	481

Figures

1.1	Map of the chapters of this book	<i>page 5</i>
1.2	Partial tree of family relationships among some Indo-European languages	12
2.1	Surviving branches within the Indo-European language family	46
2.2	The Insular Celtic Hypothesis	51
2.3	The P-Celtic Hypothesis	51
2.4	The nominative-accusative case system of Latin and the ergative-absolutive case system of Basque	65
4.1	Partial tree of the Finno-Ugric language family	112
4.2	The Finno-Ugric language family	117
4.3	The Turkic language family	124
5.1	The Northwest Caucasian language family	153
5.2	The Northeast Caucasian language family	158
5.3	The Mxedruli alphabet	164
6.1	The Semitic language family	194
7.1	The Nilo-Saharan language family (according to Greenberg 1963)	220
7.2	The Niger-Congo language family	225
8.1	The cycle of morphological type change	277
12.1	Pedersen's Nostratic proposal	416
12.2	Illyč-Svityč and Dolgopolsky's Nostratic proposal	417
12.3	Greenberg's Eurasiatic proposal	418
12.4	Nostratic vs. Eurasiatic	419
12.5	Eurasiatic as a branch of Nostratic	419
12.6	Starostin's Dene-Caucasian hypothesis	422
12.7	The Sino-Austic hypothesis	422

Tables

1.1	Cognates in English and other Germanic languages	page 13
1.2	Cognate numerals 1 through 10 in Eurasian languages	18
1.3	Some cognates in English, Latin, Greek, and Irish	18
1.4	Some cognates in Maori, Tongan, Samoan, and Hawaiian	19
1.5	Case forms of the word for ‘wolf’ in several Indo-European languages	22
2.1	Cognates from Albanian, Greek, Armenian, and other Indo-European languages	52
2.2	Cognates from Baltic, Slavic, and other Indo-European languages	53
2.3	Cognates from three branches of Slavic languages	53
3.1	Cognates in modern Indo-Aryan languages	85
3.2	Numbers in Dravidian languages and Bengali	89
4.1	Central European numerals	111
4.2	Some Finno-Ugric cognates	111
4.3	Additional Hungarian/Finnish cognates	113
4.4	Ugric numerals	117
4.5	Cognates in several Turkic languages	125
5.1	Location and direction case markers in Avar	162
5.2	Location and direction case markers in Lezgian	162
5.3	Case forms of the noun <i>kali</i> ‘woman’ in Georgian	167
5.4	Subject and object agreement markers in Georgian	168
5.5	Cognates in Ossetian and other Indo-European languages	170
5.6	Talysh and Persian cognates	171
6.1	Cognate consonantal roots in Afroasiatic	188
6.2	Prefixal verb conjugation in Afroasiatic languages	189
6.3	Cognates in Semitic languages	195
6.4	Person/number preterite forms of the regular Spanish <i>-ar</i> verb <i>hablar</i> (‘to speak’)	198
6.5	Person/number/gender past-tense forms of the Modern Hebrew verb $\sqrt{\text{K-T-V}}$ (‘to write’)	198
6.6	Person/number/gender past-tense forms of the Modern Standard Arabic verb $\sqrt{\text{K-T-B}}$ (‘to write’)	199
7.1	Noun classes in Swahili	231
7.2	Symbols for click sounds in Khoisan	236
8.1	Tones in Mandarin	257
8.2	Tones in Vietnamese	264

8.3	Tones in Thai	267
9.1	Cognates in Western and Eastern Malayo-Polynesian	293
9.2	Numerals in Austronesian	295
9.3	Some Polynesian cognates	303
9.4	Some cognates in Malagasy and Ma'anyan	317
10.1	Noun classes in Dyirbal	350
12.1	Proposed cognates in Altaic	410
12.2	Proposed cognates between Altaic, Korean, and Japanese	412
12.3	Proposed Nostratic cognates	417

Textboxes

1.1	Knights, Riders, and False Friends	15
1.2	Shared Innovations Are More Important Than Shared Retentions	20
1.3	Being a Small Language: Does Size Matter?	24
2.1	Word Order in the World's Languages	48
2.2	The Balkan Sprachbund	54
2.3	The Indo-European Homeland: Alternative Theories	58
2.4	What Do Geneticists Say?	59
3.1	A New Language Discovered in India	92
3.2	Vowel Classification	93
3.3	Which Way Are You Headed?	101
4.1	Vowel Harmony	114
4.2	“Language Nest” Program to Nurture the Enets Language	120
4.3	Agglutinative Morphology	126
4.4	Why Is the Itelmen Language Disappearing?	133
5.1	Caucasus Linguistic Expeditions	150
5.2	Ejectives	155
6.1	Buried Treasures	190
7.1	Unusual Sounds of African Languages	225
7.2	Hadza and Sandawe	235
7.3	Sound Inventories and the “Out-Of-Africa” Theory	237
7.4	Indigenous African Writing Systems	242
8.1	Hmong-Mien Languages	269
8.2	Kusunda, a Language Like No Other?	274
9.1	Pioneers of the Pacific	289
9.2	Hanunó'o Color Categories	299
10.1	Language Diversity and Density	327
10.2	In-Law Language and Other Taboos	334
10.3	Light Warlpiri: A New “Mixed” Language Discovered in Northern Australia	345
11.1	Saving Native American Languages	379
11.2	Reconstructing the Lifestyles of Three Prehistoric Amazonian Tribes	389
11.3	Unusual Number Systems	396
12.1	The Borean Proposal	423
12.2	Gestures in Oral Languages	427
12.3	Adamorobe Sign Language	432

“Focus on” Topics

1.7	Language Maps	32
2.4	Endangered Languages of Europe	65
3.4	Universals and the Parametric Theory of Language	95
4.5	Evidential Markers	136
5.5	Head-marking vs. Dependent-Marking	172
6.4	Language Contact	206
7.4	Official Languages, Trade Languages, and Creole Languages in Sub-Saharan Africa	240
8.5	Isolating Morphology and Language Change	275
9.4	The Mystery of Malagasy	311
10.3	Is There Such a Thing as a Primitive Language?	347
11.4	The Pirahã Controversy	394

Preface to the Third Edition

Ancient Greeks were among the first ones to make the observation – still relevant today – that languages differ from one another. Languages use different sounds to make up different words and combine words into phrases and sentences in very different ways. When I was growing up in Saint Petersburg, Russia, my family spoke Russian. In this language, the same word is used for ‘house’ and ‘home,’ while two different words describe what Russian speakers perceive as two different colors (*sinij* and *goluboj*), which English speakers group as ‘blue.’ Russian has soft (palatalized) consonants but lacks the English “th-sounds” (θ and ð, in a linguistic notation). English distinguishes *a blue house* from *the blue house*, while Russian has no counterpart to English articles *a* or *the*. Unlike English, Russian stacks all question words at the beginning of a sentence, as in *Kto chto gde kogda kupil?* (literally ‘Who what where when bought?’). The list goes on, and each language presents its own peculiar twists. Yet, as you will discover in this book, there are many ways in which languages are similar and use the same patterns in widely dispersed parts of the world.

This book combines geolinguistics (what languages are spoken where and by whom) with typology (in what ways languages differ and in what ways they are similar). It is organized geographically (you can see the breakdown of the world into chapters in Figure 1.1): as you travel through the world’s regions, you will learn not only what languages you might encounter there (this information is easily obtainable online) but how these languages are related to one another, what peculiarities they share, and what unique features they exhibit. You will also gain an appreciation of what is “exotic” in languages, not from the perspective of English alone, but from a more global point of view. You may even appreciate how exotic English itself is, when compared to other languages of the world.

While intended primarily as an undergraduate textbook, and complete with pedagogical features described below, this book can serve as an introduction to the world of language and the world’s languages for a general reader as well. No knowledge of linguistics or of any particular language (besides English, of course!) is presupposed, and all technical terminology is explained in the text and in the glossary at the end of the book. Numerous examples are included throughout the book to illustrate the linguistic phenomena being discussed, as are URLs of websites where you can hear the sounds of various languages, explore their grammatical properties, or investigate their geographical

distribution. (The links are also posted on the accompanying website, hosted by Cambridge University Press.)

Whether or not you are using this book for a college class, I highly recommend that you roll up your sleeves and work through at least some of the “Do It Yourself” sections at the end of each chapter. Linguistics is a little like ballet: you cannot really learn it properly without doing it yourself. All the assignments included in this book are designed to be approachable – and solvable! – with just the information provided here or in the websites that are incorporated into specific assignments. (This third edition has been enriched by the addition of many new “Do It Yourself” problems. Answers to the problems are available for course instructors only.)

Each chapter also includes a number of features whose goal is to add interesting information, ease the reading, or keep you, the reader, on your toes. Each chapter opens with a short narrative introduction, which includes a textbox listing all the language families spoken in the region, with the numbers of chapters/sections where these languages are discussed. Each chapter includes one or more longer textboxes that cover tangential topics, like indigenous African writing systems and taboo languages, and several short “Did you know?” boxes that present tidbits and factoids about the material discussed in the main text. Each chapter also includes a “Fascinating Language” sidebar that provides a very brief overview of one of the corresponding region’s languages. Some sections are partitioned from the rest of the text as “Advanced”; these segments, concerned with slightly more complicated phenomena, are meant to still be readable without any prerequisite knowledge, but you will not hurt my feelings if you decide to skip them. In addition to being updated and revised, this third edition includes several entirely new sections, including ones discussing sign languages, constructed languages, language maps, and more.

Happy exploring!

Acknowledgments

We all stand on the shoulders of giants, and this book is no exception. It is my pleasure to express my gratitude to all those whose work has helped make this book a reality.

First and foremost, I would like to thank my linguistics teachers who taught me the wonders of linguistic diversity: Mark Baker, Lisa Travis, the late Edit Doron, and especially the late Yehuda Falk – many examples from a variety of languages used throughout this book are drawn from Yehuda’s extensive hand-outs and my notes from his lectures and seminars.

I am greatly indebted to my colleagues, Stéphane Goyette, Elena “Helen” Koulidobrova, and Martin W. Lewis, who read parts or all of the manuscript, asked challenging questions, caught factual and stylistic errors, and made useful suggestions for improvement.

I am also grateful to Faruk Akkuş, David Erschler, Bert Vaux, and Ruth Kramer for suggesting literature about Berber and Cushitic languages; to Fannie Goldman and Slava Suris for inspiring me to write a much needed section about sign languages; to Maria Esipova for fascinating work on gestures accompanying speech in oral languages; and to my colleagues who shared their stories about working with lesser-studied languages, language documentation, and revitalization efforts around the world (in alphabetical order): Petr Arkadiev, Marina Chumakina, Heidi Harley, Paulina Hurwitz, Rob Munro, Charles Nydorf, Rob Pensalfini, Andrey Shluinsky, and Ghil’ad Zuckermann.

I am also obliged to the anonymous reviewers for Cambridge University Press, who provided numerous suggestions for significant improvements on the first and second editions; to the Press staff, especially Helen Barton, who has made this book (in all its three editions!) possible and who worked with me on making this third edition as effective a tool for exploration and learning as possible; and to the students in the various instantiations of my “Languages of the World” course: at NYI Summer School in Saint Petersburg, LOT Summer School in Utrecht, Stanford’s Linguistics Department, Stanford Continuing Studies Program, Santa Clara University OLLI, San Francisco State University OLLI, the Forum at Rancho San Antonio, and the Nueva High School, all of whom helped me develop the material for this book.

Last but not least, I am most grateful to Kirill Glebovitskiy for his love and support.

Abbreviations Used in the Glosses

1	first person	DU	dual
2	second person	DUR	durative
3	third person	DYN	dynamic tense
A	agent-like argument of a canonical transitive verb	ERG	ergative
ABL	ablative	EVID	evidential
ABS	absolutive	EXC	exclamatory
ACC	accusative	F	feminine
AD	additive	FACT	factive
ADESS	adessive	FIN	finite
ADV	adverbial clause marker (complementizer)	FOC	focus
ALL	allative case	FORM	formal
ANTIP	antipassive	FUT	future
AOR	aorist	GEN	genitive
APPL	applicative	HAB	habitual
ART	article	HON	honorific
ASP	aspect	INDEF	indefinite
AT	Actor-topic	INDIC	indicative mood
AUX	auxiliary	INDIRECT	indirect case
BT	Beneficiary-topic	INS	instrumental
CAUS	causative	INTNS	intensifier
CLF	classifier	IPFV	imperfective
CMPL	completive	IT	Instrument-topic
COMP	complementizer	LOC	locative
COND	conditional	M	masculine
CONT	continuative	MID	middle voice
CS	construct state	NEG	negation
CT	Circumstantial-topic	NFUT	non-future
DAT	dative	NMLZ	nominalizer/nominalization
DCL	declarative	NOM	nominative
DEF	definite	NSG	non-singular
DEM	demonstrative	OBJ	object
DIRECT	direct case	OBL	oblique
DT	direction-topic	P	patient-like argument of a canonical transitive verb
		PART	particle

PASS	passive	REL	relative
PFV	perfective	REPET	repetitive
PL	plural	SBJ	subject
POSS	possessive	SBJNCT	subjunctive
POTENT	potential	SG	singular
PROG	progressive	SGL	singulative
PRS	present	SIM	simple tense
PRTV	partitive	TAM	tense-aspect-modality
PST	past	TOP	topic
PTCPL	participle	TR	transitive
PUNC	punctual	TT	Theme-topic
Q	question particle/marker	UNPOSS	unpossessed noun
REFL	reflexive	VOC	vocative