The Cambridge Handbook of

Linguistic Typology

edited by Alexandra Y. Aikhenvald and R. M. W. Dixon

The Cambridge Handbook of Linguistic Typology

Linguistic typology identifies both how languages vary and what they all have in common. This handbook provides a state-of-the art survey of the aims and methods of linguistic typology, and the conclusions we can draw from them. Part I covers phonological typology, morphological typology, sociolinguistic typology and the relationships between typology, historical linguistics and grammaticalization. It also addresses typological features of mixed languages, creole languages, sign languages and secret languages. Part II features contributions on the typology of morphological processes, noun categorization devices, negation, frustrative modality, logophoricity, switch reference and motion events.

Finally, Part III focuses on typological profiles of the mainland Southeast Asia area, Australia, Quechuan and Aymaran, Eskimo-Aleut, Athabaskan (Dene), Iroquoian, the Kampa subgroup of Arawak, Omotic, Semitic, Dravidian, the Oceanic subgroup of Austronesian, and the Awuyu-Ndumut family (in West Papua). Uniting the expertise of a stellar selection of scholars, this Handbook highlights linguistic typology as a major discipline within the field of linguistics.

ALEXANDRA Y. AIKHENVALD is Australian Laureate Fellow, Distinguished Professor, and Director of the Language and Culture Research Centre James Cook University, Australia. She is an expert on languages and cultures of Amazonia and the Sepik region of Papua New Guinea, in addition to linguistic typology, general linguistics, language contact and several other areas.

R. M. W. DIXON is Professor and Deputy Director of the Language and Culture Research Centre at James Cook University, and a Fellow of the British Academy. He has published grammars of a number of Australian languages, a dialect of Fijian and of Jarawara, an Amazonian language. He is an expert in linguistic typology and the author of a classic three-volume work on basic linguistic theory.

CAMBRIDGE HANDBOOKS IN LANGUAGE AND LINGUISTICS

Genuinely broad in scope, each handbook in this series provides a complete state-of-the-field overview of a major sub-discipline within language study and research. Grouped into broad thematic areas, the chapters in each volume encompass the most important issues and topics within each subject, offering a coherent picture of the latest theories and findings. Together, the volumes will build into an integrated overview of the discipline in its entirety.

Published titles

The Cambridge Handbook of Phonology, edited by Paul de Lacy
The Cambridge Handbook of Linguistic Code-switching, edited by Barbara E. Bullock and Almeida Jacqueline Toribio
The Cambridge Handbook of Child Language, Second Edition, edited by Edith L. Bavin and Letitia Naigles
The Cambridge Handbook of Endangered Languages, edited by Peter K. Austin and Julia Sallabank
The Cambridge Handbook of Sociolinguistics, edited by Rajend Mesthrie
The Cambridge Handbook of Pragmatics, edited by Keith Allan and Kasia M. Jaszczolt
The Cambridge Handbook of Language Policy, edited by Bernard Spolsky
The Cambridge Handbook of Second Language Acquisition, edited by Julia Herschensohn and Martha Young-Scholten
The Cambridge Handbook of Biolinguistics, edited by Cedric Boeckx and Kleanthes K. Grohmann
The Cambridge Handbook of Generative Syntax, edited by Marcel den Dikken
The Cambridge Handbook of Communication Disorders, edited by Louise Cummings
The Cambridge Handbook of Stylistics, edited by Peter Stockwell and Sara Whiteley
The Cambridge Handbook of Linguistic Anthropology, edited by N. J. Enfield, Paul Kockelman and Jack Sidnell
The Cambridge Handbook of English Corpus Linguistics, edited by Douglas Biber and Randi Reppen
The Cambridge Handbook of Bilingual Processing, edited by John W. Schwieter
The Cambridge Handbook of Learner Corpus Research, edited by Sylviane Granger, Gaëtanelle Gilquin and Fanny Meunier
The Cambridge Handbook of Linguistic Multicompetence, edited by Li Wei and Vivian Cook
The Cambridge Handbook of English Historical Linguistics, edited by Merja Kytö and Päivi Pahta
The Cambridge Handbook of Formal Semantics, edited by Maria Aloni and Paul Dekker
The Cambridge Handbook of Morphology, edited by Andrew Hippisley and Greg Stump
The Cambridge Handbook of Historical Syntax, edited by Adam Ledgeway and Ian Roberts
The Cambridge Handbook of Linguistic Typology, edited by Alexandra Y. Aikhenvald and R. M. W. Dixon

Forthcoming

The Cambridge Handbook of Areal Linguistics, edited by Raymond Hickey
The Cambridge Handbook of Cognitive Linguistics, edited by Barbara Dancygier

The Cambridge Handbook of Linguistic Typology

Edited by

Alexandra Y. Aikhenvald and R. M. W. Dixon

Language and Culture Research Centre, James Cook University

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE
UNIVERSITY PRESS

University Printing House, Cambridge CB2 8BS, United Kingdom

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/9781107091955

First published 2017

Printed in the United Kingdom by Clays, St Ives plc

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

Library of Congress Cataloguing in Publication data
Aikhenvald, A. Y. (Aleksandra Yurievna) editor. | Dixon, Robert M. W., editor.
The Cambridge handbook of linguistic typology / edited by Alexandra Y. Aikhenvald and R. M. W. Dixon.
Cambridge : Cambridge University Press, [2016] | Includes index.
LCCN 2016014769 | ISBN 9781107091955
LCSH: Typology (Linguistics) – Handbooks, manuals, etc.
LCC P204 .C263 2016 | DDC 410.1–dc23
LC record available at https://lccn.loc.gov/2016014769

ISBN 978-1-107-09195-5 Hardback

Contents

Figures and Diagrams

Diagrams

Figures

Tables

Maps

Contributors

Willem F. H. Adelaar, Professor, Leiden University Centre for Linguistics, Leiden University, W.F.H.Adelaar@hum.leidenuniv.nl
Alexandra Y. Aikhenvald, Distinguished Professor and Director, Language and Culture Research Centre, James Cook University, alexandra.aikhenvald@jcu.edu.au
Felix K. Ameka, Professor, Leiden University Centre for Linguistics, F.K.Ameka@hum.leidenuniv.nl
Azeb Amha, Researcher, Africa Studies Centre, Leiden University, A.Amha@asc.leidenuniv.nl
Peter Bakker, Professor, University of Aarhus, linpb@cc.au.dk
David Beck, Professor, University of Alberta, dbeck@ualberta.ca
Aymeric Daval-Markussen, PhD scholar, University of Aarhus, linadm@cc.au.dk
R. M. W. Dixon, Professor and Deputy Director, Language and Culture Research Centre, James Cook University, Robert.Dixon@jcu.edu.au
N. J. Enfield, Professor, The University of Sydney, nick.enfield@sydney.edu.au
Michael Fortescue, Emeritus Professor, University of Copenhagen, fortesq@hum.ku.dk
Valérie Guérin, Adjunct Research Fellow, Language and Culture Research Centre, James Cook University, Valerie.Guerin@jcu.edu.au
Harry van der Hulst, Professor, University of Connecticut, harry.van.der.hulst@uconn.edu
Silvia Luraghi, Professor, Università degli Studi di Pavia, luraghi@unipv.it
Matti Miestamo, Professor, General Linguistics, University of Helsinki, matti.miestamo@helsinki.fi
Elena Mihas, PostDoctoral Research Associate, Language and Culture Research Centre, James Cook University, Elena.Mihas@jcu.edu.au

Marianne Mithun, Professor, University of California at Santa Barbara, mithun@linguistics.ucsb.edu
Edith Moravcsik, Professor Emerita, University of Wisconsin-Milwaukee, edith@uwm.edu
Heiko Narrog, Professor, Tohoku University, narrog@gmail.com
Simon E. Overall, PostDoctoral Research Fellow, Language and Culture Research Centre, James Cook University, Simon.Overall@jcu.edu.au
Nick Palfreyman, Research Fellow, International Institute for Sign Languages and Deaf Studies (iSLanDS), University of Central Lancashire, nickpalfreyman@cantab.net
Thomas E. Payne, Research Associate, University of Oregon and Senior Linguistics Consultant, SIL International, tpayne@uoregon.edu
Eric Pederson, Professor, University of Oregon, epederso@uoregon.edu
Willem de Reuse, Research Professor, University of North Texas, WillemDeReuse@my.unt.edu
Keren Rice, Professor, University of Toronto, rice@chass.utoronto.ca
John R. Roberts, Senior Linguistics Consultant, SIL International, dr_john_roberts@sil.org
Aaron D. Rubin, Professor, Penn State University, arubin@psu.edu
Sanford Steever, Independent scholar, New Canaan, Connecticut, sbsteever@yahoo.com
Anne Storch, Professor, Institut für Afrikanistik, University of Cologne, astorch@uni-koeln.de
Peter Trudgill, Professor, University of Agder, peter.trudgill@unifr.ch
Lourens de Vries, Professor, Free University of Amsterdam, l.j.de.vries@vu.nl
Ulrike Zeshan, Professor and Director of the International Institute for Sign Languages and Deaf Studies (iSLanDS), University of Central Lancashire, uzeshan@uclan.ac.uk

Acknowledgements

We owe the idea of putting this book together to Helen Barton, at Cambridge University Press. Without her unfailing support and assistance at every stage, this project would not have been possible. We are grateful to all the contributors to this volume, for their chapters and for the comments we received. Many thanks go to Professor Nola Alloway, Dean of the College of Arts, Society and Education at James Cook University for her moral and financial support. We owe a special debt of gratitude to Brigitta Flick and Jolene Overall, for meticulous editorial assistance, checking the manuscripts and formatting them. A very big thank you goes to Amanda Parsonage for her assistance and support at the Language and Culture Research Centre.

Abbreviations

1	first person
1 > 2, 2 > 3, etc.	first person acts on second person, second person acts on third person, etc.
2	second person
3	third person
3.3′.CJ	third person subject with third person object conjunct order verb
4	first person inclusive or fourth person, impersonal pronoun
A	transitive subject function; Actor macrorole; Actor
ABE	abessive
ABL	ablative
ABS	absolutive
ACC	accusative
ACT	active
AD	addressee deictic
ADD	additive
ADEL	adelative
ADJ	adjective
ADJZ	adjectivizer
ADV	adverb
ADVST	adversative
AFF	affirmative
AFOC	argument focus marker
AG	agentive
AGAIN	repetitive aspect, topic shift
AGR	agreement
AGT	agent
AL	alienable

ALL	allative
AN	Austronesian
AN	animate
ANA	action narrowly averted
ANAPH	anaphoric
AND	andative non-singular, singular transitive
ANIM.INTR	animate intransitive
ANP	adnominal verb
ANT	anterior
ANTIC	anticipatory
ANTICAUS	anticausative
ANTIP	antipassive
AntSu	anticipatory subject
AOR	aorist
APPL.APPLIC	applicative
APPL.GEN	general applicative
APPL.PRES	presential applicative
AREA	areal
APR	apprehensive
ARG	argument
ART	article
ASS.P	associative plural
ASP	aspect
ASS	assertive
ASSOC	associative
ATR	Advanced Tongue Root
ATTN	attenuative
aug	augmented
AUX	auxiliary
Aym	Aymaran
BEN	benefactive
C	common gender; consonant; controlling clause
C.EXIST	ceased existence
CAUS	causative
CIRC	circumstantial
CIS	cislocative
CJ	conjugation
CL	class
CLF	classifier
CLT	clitic
CMPL	completive
CND	conditional
CNG	connegative
CNJ	conjunctive verb
CNSTR	construct

CNTREXP	counterexpectation
CNTR.FACT	counterfactual condition
CNV	converb
COMIT	comitative affix
COMP	complementiser
COMPL	completive
COND	conditional
CONJ	conjectural (evidential)/conjunction
CONN	connective
CONSTR	construct
CONT	contemporative
CONTIN	continuative
CONTR	contrast
CONV	converb
COORD	coordination
COP	copula
CT	circumstantial topic; class term
CUST	customary
CVB	converb
D	dependent
D.CLF	d-classifier
DAT	dative
DEC, DECL	declarative
DEF	definitive; definite
DEIC	deictic particle
DEM	demonstrative
DepSu	dependent clause subject
DESID	desiderative
DET	determiner
DIM	diminutive
DIR	directive; directional
DIST	distal
DISTR	distributive
DL	dual
DNA	Direct Negation Avoidance
DO	direct object
DOM	differential object marking
DR	different reference
DS	different subject
DSTL	distal demonstrative
DU	dual
DUR	durative aspect
DV	duplicative
E	epenthetic element
EMP, EMPH	emphatic

EP	epenthetic
ERG	ergative
ES	echo subject
EV	direct evidential
EVI	evidential
EX	existential
EXC, exc, EXCL	exclusive
EXH.FOC	exhaustive focus
EXIST	existential
EXP	experienced
F, f	feminine
FACT	factitive
FAR.PST	far past
FEM	feminine
FI	feminine/indefinite
Fj/J	the ratio of fusional morpheme junctures to all morpheme junctures
FOC	focus
FOOT	'foot' as body part prefix
FP	free pronoun
FREQ	frequentative
fric	fricative
FRUST	frustrative
FS	factual stem
FUT	future
GEN	genitive
GENZ	generalized
GER	gerund
H	high tone or pitch
HAB	habitual
HEDGE	hedging device
HL	Highlighter
HN	head noun
HON	honorific
HRSY	hearsay
HUM	human
IDPH	ideophone
IE	Indo-European
IMMED	immediate
IMP	imperative
IMPERS	impersonal
IMPF	imperfective
IMPL	implicated
IMPST	immediate past
INAL	inalienable

INAN	inanimate
INC, inc	inclusive
INCEP	inceptive
INCL	inclusive
INCOMPL	incompletive
IND	indicative
INDEF	indefinite
INDEP	independent
INESS	inessive
INF	infinitive
INFL	inflection
INFR	inferred evidential
INFRN	inferential
INGR	ingressive
INST	instrumental
INSV	inessive
INT	intensifier
INTENT	intentional
INTER	interrogative
INTERJ	interjection
INTR	intransitive
INTS	intensive
INTSINT	International Transciption System for Intonation
INV	inverse
IPFV	imperfective
IRR	irrealis
ITER	iterative
ITG	intangible
IVC	impersonal verb construction
KBo	Keilschrifttexte aus Boghazköi
L	low tone or pitch
LF	low fall
LIG	ligature
LIM	limitative case
LK	linker
LOC	locative
LOC.CL	locative classifier
LOG	logophoric pronoun
LogA	speaker logophoric pronoun
LogB	addressee logophoric pronoun
LP	linking particle
M, m	masculine
M	mid tone

MAL	malefactive
MASC	masculine
MED	medial (function similar to that of 'converb')
MID	middle
min	minimal
MIR	mirative
MOD	modifier
M.PL.AGT	masculine plural agent
M/W	the number of morphemes per grammatical word
N	neuter gender; noun
N.AGT	non-agentive
NARR	narrative marker
NCL	noun class
NEG	negative
NENA	Northeastern Neo-Aramaic
NEWSIT	new situation
NF	non-finite; non-feminine
NFUT	non-future
NFS	non-factual stem
NH	non-human
NHYP	non-hypothetical
NM, nmasc	non-masculine
n-min	non-minimal
NMLZ	nominalizer
NMR	non-macrorole argument
NMZ	nominalizer
NOM	nominative
NOML	nominal
NON-EXP	non-experienced
NONFUT	non-future
NONPUNCT	nonpunctual
NON3	non-third person
NON_SBJ	non-subject
NOUN.CL	noun classifier
NP	noun phrase
NPST	nonpast
NREF	non-referential
NS	non-singular
NSG, nsg, n-sg	non-singular
NTS	non-topic subject
NUM.CL	numeral classifier
O	object
OBJ	object
OBL	oblique
OBV	obviative

OPT	optative
OV	object-verb order
P	patient-like
P2	second position
PAA	Proto-Afroasiatic
PAR	partitive
PART	participle
PART.SG	partitive singular
PASS	passive
PAT	patient
PD	Proto-Dravidian
PERF	perfective aspect
PERS	person
PERT	pertensive
PF	perfect
PFV	perfective
PI	polar interrogative
PIE	Proto-Indo-European
PK	personal knowledge
PL, pl, plur, PLUR	plural
PNG	person/number/gender
POc	Proto-Oceanic
POL	polite
POS	positive
POSS	possessive; possession
POSS.CL	possessive classifier
POSTP	postpositional
POSTP:PURP	purposive postpositional
POT	potential
PP	past perfect
PP, Pp	postposition; positive polarity
PR	present
PR.PART	pragmatic particle
Pr	preposition
PRED	predicate
PREP	preposition
PRES	present
PRF	prefix
PRO	pronoun
PROG	progressive
PROX	proximal
PRS.PRF	present perfect tense
PRT	preterit
PS	past
PSA	privileged syntactic argument

PSR	possessor
PST	past
PTC	particle
PTCP	participial
Q	question element
QT	quotative marker
Qu	Quechuan
QUAL	qualifier
QUOT	quotative
R	realis; also recipient
RCP	reciprocal
RDP	reduplicant
RED	reduplicative
REDUPL	reduplication
REAL	realis
REC.PST	recent past
REFL	reflexive
REGR	regressive
REL	relative; relative clause marker
REL.CL	relational classifier
REM	remote
REMP	remote past tense
REP	reportative (evidential)
RES	resultative
REV	revisionary
RLS	realis
RP	referential phrase; also reporting pronoun
RPT	repetitive
RRG	Role and Reference Grammar
RS	relativized subject
RV	Rigveda
S	intransitive subject; subject; also singular
SAP	Speech Act Participant
SBJ	subject
SBJV	subjunctive
SC	stem class
SEP	separative
SEQ	sequential
SF	singular feminine
SG/sg	singular
SGLT	singulative
Sh	short
SH	superhigh tone or pitch
SIM/SIMUL	simultaneous
Skr/Skt	Sanskrit

SM	singular masculine
SMI	semeliterative
SMR	same reference
SN	singular neuter
SO	same object
SOC	sociative
SOV	subject-object-verb order
SP	specifier
SQ	sequential
SR	switch reference
SS	same subject
StBoT	Studien zu den Bogazköy-Texten
ST.PL	stative plural
STAT	stative
STR	strong
StressTyp	stress typology database
SU, SUB, SUBJ	subject
SUBJV	subjunctive
SUBORD	subordinate
SUFF	suffix
SVO	subject verb order
SylTyp	syllable typology database
T	theme
TI2	transitive inanimate verb type 2
TA	transitive animate verb
TAM	tense and/or aspect and/or modality and/or mood
TEL	telic, contrastive emphasis
TEMP	temporal
TERM	terminative
THM	thematic
ToBI	tone and break indices
TODP	today's past tense
TOP	topic
TOT	totalitative
TR, TRANS	transitive
TRC	transitive controlling clause
TRM	terminative derivational suffix
TR.2	type 2 transitive verb
TV	terminal vowel
U	undergoer macrorole
ua	unit augmented
UNSP	unspecified
UPSID	Universal Phonological Segment Inventory Database
V, VB	verb
VD	voiced

VBZ	verbalizer
vcl	voiceless
VCL	verbal classifier; also verb closing element
VIS	visual evidential
VL	voiceless
VN	verbal noun
VO	verb-object order
VOC	vocative
VOT	voice onset time
vow	vowel
VP	verb phrase
VSO	verb-subject order
vv	voice/valence
Xtone	cross-linguistic tonal database

Glosses for Sign Languages

SIGN	gloss for a manual sign
SIGN-SIGN	hyphenated gloss for a single sign that needs more than one word to express its meaning
SIGN–	sign held in its final position for a length of time
IX or INDEX	pointing signs, followed by further information about the reference of the point
1,2	subscripts indicating first and second person reference
SIGN#SIGN	simultaneous morphology
SIGN^SIGN	host-clitic combination
_____t	non-manual marking for topicalisation
_____y/n	non-manual marking for polar question
_____low-br	lowered eyebrows
_____brow-raise	raised eyebrows
_____neg-tilt	backward head tilt for negation
|	clause boundary

Further Conventions

=	clitic break
–	affix boundary
≈	varies with

1

Introduction: Linguistic Typology – Setting the Scene

Alexandra Y. Aikhenvald and R. M. W. Dixon

Linguistic typology as a discipline is all-embracing. Its subject matter is rigorous scientific investigation of cross-linguistic variation in every domain of language, and thus of the limits and of the possibilities of human language. Typology investigates kinds of linguistic phenomena within a language, and across languages. It may also delineate types of languages and classify them.

Typologists work out constraints on language structures and formulate predictions as to what one could expect in a language. Typology is central for inductively based cross-linguistic generalizations. And it is a testing ground for any theoretical statement about what a language, or a language feature, may be like. Typological generalizations – based on good-quality reliable sources – help predict what is more likely, and what is less likely, to occur in a language. Ideally, typology goes hand in hand with explanation – why the language is the way it is. In this introductory chapter, we address linguistic typology in its various guises, offering a brief glimpse into typological parameters, the scope of linguistic typology and its relationship with linguistic analysis, historical linguistics and language contact. The last section presents an overview of this volume.

1.1 Linguistic Typology in Its Various Guises

Linguistic typology involves a number of interrelated areas of intellectual pursuit.

FIRSTLY, typology is a comprehensive investigation of cross-linguistic variation. A major premise of linguistic typology is, in Comrie's (1990: 139) words, 'that different languages vary in a number of different respects', and, 'more importantly, that this cross-linguistic variation is amenable to

systematic study'.[1] It aims at uncovering, and understanding, the range of structures and parameters available in human languages for expressing a wide variety of core concepts.

SECONDLY, typology involves establishing cross-linguistic generalizations that hold for all, or most, human languages, and correlations between linguistic phenomena. These generalizations are often called 'language universals'.

Language universals can be absolute. A statement that every spoken language has vowels and consonants holds everywhere and is an 'absolute' universal. So does a statement that every language has a way of asking a question, forming a command, or articulating negation.

Universals also reflect tendencies which are well attested cross-linguistically. A statement that languages tend to have more oral than nasal vowels reflects such an overall tendency. A universal tendency can be phrased as an implicature. An example of an 'implicational universal' is a statement that if a language has a high front nasal vowel, it will also have a low front nasal vowel.[2] Implicational universals may allow us to understand the reason 'why' language is the way it is, and even predict the ways in which it might develop. Absolute universals – which have no exceptions whatsoever – are few. Most statements of general value are cast in the form of universal tendencies which allow some exceptions, but hold overall. For instance, Greenberg's (1963: 95) Universal 45 states 'If there are gender distinctions in the plural of the pronoun, there are some gender distinctions in the singular also.' There are just a few exceptions to this statement – for instance, independent personal pronouns in Tamachek, a Tuareg Berber language, distinguish masculine and feminine genders in all three persons in the plural, but only in the second person in the singular.[3]

Linguistic typology helps establish correlations between subparts of languages. A choice made in one part of a linguistic system may limit a set of options in some other area. For example, if a language has multi-word discontinuous serial verb constructions, it is likely to be analytic (rather than synthetic) in its morphological structure. If a language has an overt marker for plural number on nouns with inanimate reference (such as 'tree' or 'house'), it is pretty much bound to also mark number on nouns with animate and human reference (such as 'dog' or 'man').[4]

Finally, linguistic typology aims at offering explanations – why there are certain linguistic universals and universal tendencies, and what their limits may be.[5] There is here a direct link to how humans communicate and perceive the world around them, and how languages may reflect physical, social and cultural environment.

Linguistic typology is often viewed – in a somewhat restrictive way – as a classification of languages by their synchronic grammatical features, 'according to their general structure rather than according to their historical or geographical relationship' (Bazell 1958: 3). This is congruent with

a general non-specifically linguistic understanding of the term 'typology', as 'the study of classes with common characteristics; classification' (as given in *The Oxford English dictionary online*). Typological classifications have indeed gained some currency, and features and parameters employed in such procedures have been helpful in understanding how human languages work. This is what we turn to now.

1.2 Typological Parameters and the Scope of Typology

A typology of languages can aim at encompassing the totality of linguistic structure under one parameter. Or it can deal with one domain of language structure. We now turn to the scope of linguistic typology, and typological parameters.

1.2.1 Typological Classification and Typological Parameters

A typological parameter may aim at classifying a language as belonging to a general type.[6] Earlier typological endeavours aimed at a classification of languages into morphological types based on the number of morphemes per word and on the techniques of combining morphemes into words.

In the first instance, languages were classified depending on whether words neatly divide into meaningful parts (morphemes) with one meaning each (further discussion is in Payne, Chapter 3 of this volume). This involves three ideal types:

i. ISOLATING languages, like Vietnamese and Mandarin Chinese, where every form has one meaning.
ii. AGGLUTINATING languages, like Turkish, where a form may consist of several morphemes but the boundaries between them are clear-cut.
iii. FUSIONAL languages, like Latin, where one form combines many meanings and is not easily segmentable.

An agglutinating language typically has a one-to-one correspondence between a morpheme and its meaning. A morpheme tends to have an invariant shape which makes it easy to identify. For instance, in Turkish *ev-ler-im-de* (house-PLURAL-1sg-LOCATIVE) 'in my houses' every morpheme corresponds to a string of one or several phonemes. In a FUSIONAL language, like Latin, morphemes which have to be distinguished grammatically combine into one hard-to-parse form. For instance, *i*, the shortest form in Latin, has a complex meaning of 'you singular go! (imperative)'.

This erstwhile tripartite typological classification assumed that affixation is the only morphological process. To amend this, Sapir (1921: 126) posited a fourth type of language, where morpheme junctures are characterized by 'internal changes (reduplication, vocalic and consonantal

change, changes in quantity, stress and pitch)'. He called this language type 'symbolic' (the term, and the concept, are rarely used today; see Beck, Chapter 11 of this volume, for a detailed typological analysis of morphological processes across the world's languages). However, hardly any language fits one type exactly. There tends to be a mixture of techniques. Quite a few languages are best referred to as 'basically agglutinating with some fusion'. For instance, Tariana, an Arawak language from Brazil, allows for some change of form on prefixal boundaries, but not on suffixal ones. The classificatory parameters are still in use, but they are hardly exclusive.[7]

Alternatively, languages can also be classified depending on a number of meaningful parts – that is, morphemes – within a word.

i. ANALYTIC languages tend to have a one-to-one correspondence between word and morpheme; they have few if any bound morphemes. Vietnamese and Mandarin Chinese are good examples of analytic languages.
ii. In SYNTHETIC languages, a word consists of several morphemes, and there are numerous bound morphemes. Hungarian and Latin are synthetic languages. Some languages are only mildly synthetic. In English, and in many languages of the Jê family from South America, the number of morphemes per word is often not more than two. Languages of the Eskimo-Aleut family such as Yupik and Greenlandic, and many languages of Amazonia, North America and northern Australia are at the opposite extreme: a word often contains a long string of morphemes (see Chapters 22–25).

In highly synthetic languages, bound morphemes express notions that would be expressed with lexical items in languages which are less synthetic. An example from Mohawk, a Northern Iroquoian language (Mithun, Chapter 24 of this volume, example 29), shows just how much information can be packed into one verb.

(1) Ahsani'tskwahra'tsherakarhátho'
aa-hs-[an-i'tskw-a-hr-a'tsher]-a-karhat-ho-'
IRR-2SG.AGT-[MID-rump-LINKER-set.on-NMLZ]-LINKER-turn-REVERSIVE-PFV
'You might [thing you set your rump on] tip over' = 'You might tip over a chair'

The term 'polysynthetic' describes such highly synthetic 'mega-rich' structures.[8]

The degree of synthesis and the treatment of morphological boundaries are relatively independent typological parameters. However, they do intersect. A highly analytic language will tend to be isolating, as is the case in Mandarin Chinese and Vietnamese. For instance, Turkish is synthetic and agglutinating, and Latin is synthetic and fusional. English has strong

analytic tendencies and is fusional. The two sets of parameters differ in yet another way. Each of 'isolating', 'agglutinating' and 'fusional' may be used to characterize the treatment of morphological boundaries in a language. 'Synthetic' and 'analytic', on the other hand, refer to a continuum. A language can be more, or less, synthetic or analytic.

Morphological typology based on these two sets of properties dominated linguistic typology as a major parameter for classifying a language as a whole, throughout the nineteenth and early twentieth century. Thomas Payne, in Chapter 3, offers a detailed analysis of history and approaches to morphological typology.[9]

A further parameter for classification of languages, originally put forward by Greenberg (1966), concerns the order of the meaningful elements. This originally included position of bound morphemes within a word (prefixes and suffixes), order of words within noun phrases and order of constituents within a clause or a sentence. The order of syntactic constituents within a clause varies across languages. In English, the verb has to follow the subject, as in *I saw a dog*. But in Fijian, the verb generally comes first. So, English is verb-medial, and Fijian is verb-initial, while Manambu, from New Guinea, is verb-final. Order of individual words within a constituent – especially a noun phrase – is another parameter. For instance, English and Latin have prepositions which come before the noun (as in *post hominem* and its English translation 'after (a) man'), Hungarian has postpositions which come after the noun, as in *ember után* (man after) 'after a man', and Estonian has both, as in *sõja pärast* (war:genitive.singular after/because) 'because of the war' and *pärast sõda* (after war:partitive. singular) 'after the war'.

A typology of languages based on the order of elements involves some correlations between orders within different kinds of constituents. For instance, languages with verb-initial constituent order tend to have prepositions (as does, for instance, Modern Welsh). Languages with a subject-object-verb order are likely to have postpositions; within a noun phrase, with possessor and adjective modifiers preceding the head noun.

Both morphological typology and the typology based on the order of meaningful elements attempt at classifying languages as systems in their totality. The typology of order of meaningful elements was perhaps the first attempt, in linguistic typology, to establish implicational relations between different features within a language. It demonstrated that typological generalizations may have a predictive power.[10]

A further typological parameter for an overall classification of languages was suggested by Nichols (1986). Languages are divided into head-marking and dependent-marking (verb is seen as 'head' of a clause). In a dependent-marking language, the morphological information

concerning the relationship between the head and its dependents will attach to the dependent. For example, in Japanese, a dependent-marking language, the verb is unmarked for the relations with its dependents. The dependents (e.g. nouns or pronouns expressing arguments or obliques) are marked with post-nominal 'particles' for a particular relation they have with the verb as head. This is illustrated in (2):

(2) Taroo ga Ziroo ni hon o yatta
Taro NOM Jiro DAT book ACC gave
'Taro gave a book to Jiro'

A typical head-marking language is Abkhaz, from the north-west Caucasian family. Here only the verb, as head of the clause, is marked for the person, number and noun class of its arguments. The arguments themselves are not marked:

(3) a-xàc'a a-pħ°ə̀s a-š°q°'ə̀ (ø-)lə̀-y-teyt'
the-man the-woman the-book it-to:her-he-gave
'The man gave the book to the woman'

Many languages (including Indo-European) display features of both head- and dependent-marking. The applicability of this binary division depends on the theoretical orientation of the scholars. Many have argued against extending the notion of head from a noun phrase to a clause.[11]

A typological parameter can focus on a subpart of a language. One such parameter involves the expression of prototypical grammatical relations in a clause. The universal grammatical relations within a clause are A (transitive subject), S (intransitive subject) and O (object) (see Dixon 1994, 2010a: 98–9). Some languages mark S and O in the same way (absolutive case) and A differently (ergative case). Others mark S and A in the same way (nominative case) and O differently (accusative case). Languages can thus be classified as absolutive-ergative or as nominative-accusative. And there may be other localized parameters, each of which helps investigate the limits and the possibilities of typological variation in one particular domain.

We now turn to the modes of typological investigation of linguistic phenomena.

1.2.2 What to Compare: Two Modes of Linguistic Typology

Two basic modes of linguistic typology are **intra-language typology** and **extra-language typology.** The two are intertwined.

Intra-language typology involves comparing a feature in one language with similar features in other languages, in terms of a defined set of theoretical parameters. A typological study can investigate the following, across languages:

- Structures – for instance, the make-up of a syllable, a word or other unit in phonology, or of a complex predicate, or a noun phrase, or a clause in grammar;
- Systems – for instance, system of tense-aspect choices, or number systems, in grammar, or consonant systems in phonology;
- Construction types in grammar – for instance, relative clause constructions, speech report constructions or complement clause constructions;
- Mechanisms of marking particular categories – for instance, marking of possession within a noun phrase or a clause, or syntactic relations within a clause.

Any phenomenon in language can be analysed from a typological perspective. A feature, a system or a construction can be compared across languages, so as to establish the limits of cross-linguistic variation and suggest possible interrelations with other properties of a language. In this respect, linguistic typology constitutes the core of linguistics as a discipline. A typologically informed analysis allows us to determine what is to be expected in a language, what is typologically plausible and what is unusual and warrants further study.

Comparing isolated entities without taking account of the whole system is unhelpful and uninformative – for instance, comparing the occurrence of an unrounded central close vowel phoneme /i/ without the phonological system in general, or looking at the realization of a visual evidential, across languages, without paying attention to the evidentiality system (grammaticalized marking of information source) in each language to be compared. The proper method is to compare complete vowel systems, or complete systems of evidentials. The role of /i/ within the system will vary depending on the size of the vowel system. Its phonetic realization and function in a three-vowel system will be quite different from those in a ten-vowel system. Similarly, a visual evidential in a three-term system (visual, inferred, reported) is likely to subsume a wide range of first-hand information. A visual evidential in a larger term system (for instance, visual, non-visual, inferred, assumed, reported) is likely to be restricted to just visually acquired, or 'seen', information.

The other mode of typology relies on some phenomenon in the real world, rather than a category or a feature of the abstract system of phonology and grammar which result from a linguist's analysis. This can be called **extra-language typology**. A typologist may wish to look at how directionality is encoded across the languages of the world. This may involve directional verbs, directional demonstratives or special directional markers. Or there may be specific lexical items referring to direction.

If an extra-language typology describes how something in the real world – time, commands, direction, or definiteness – can be shown through a grammatical system (such as a tense system for expressing

time, or imperative system for expressing commands), the typologist should then turn to the intra-language typology of that system. An extra-language typology is not an end in itself. If it is to have predictive and explanatory power, it must be seen as a conduit to one or several intra-language typological studies.

For instance, one can consider different kinds of possible information sources – vision, hearing, smell, touch, inference, assumption and verbal report. In order to come up with sensible generalizations concerning the expression of information source within languages, one should provide an in-depth analysis of grammatical systems which express these meanings, within languages.

An extra-language typology examines how something outside a language system is coded within that system. To have any scientific validity, it must be augmented by an intra-language typological study of the coding mechanism, and then the variations.

During the past decades, typological investigations have involved every domain of language. Common topics in phonological typology include syllable structure, systems of consonants, and of vowels and tones. Grammatical typology addresses the typology of word classes, and grammatical systems and categories – tense, aspect, evidentiality, gender and other noun categorization devices, case and grammatical relations, and negation. For each study – in a language and across languages – care should be taken to note how one category may interact with another. So, for instance, if a language has evidentiality, this will always apply in past tense and declarative mood but not necessarily in imperative mood and in future tense. The issue of dependencies between grammatical categories is fascinating for typological studies. So, there may be fewer distinctions in person, gender, number, tense, aspect and evidentiality under negation than in positive clauses (see Aikhenvald and Dixon 1998 for a preliminary discussion). The ways in which grammatical categories interrelate and depend upon each other may shed light on their role in human cognition, helping linguists predict what is likely, and what is not, in a particular language.

Lexical typology is particularly informative when it is viewed in interaction with grammar. Dixon (1982; 2004; 2010b: 62–114) demonstrated that if a language has even a small adjective class, it is likely to include lexemes from the semantic types of Dimension, Age, Colour, and Value. Lexical typologies may involve colour terminologies, body part terminologies and terminologies for kinship. Further fruitful topics may involve sets of verbs referring to motion; or to posture and directionality.

1.2.3 Implications and Prediction

Typological generalizations about any linguistic phenomena can be usefully stated as implications 'if x, then y', also referred to as

implicational hierarchies.[12] Implicational hierarchies consist of a chain of cross-linguistically confirmed generalizations which are dependent upon each other. Such a hierarchy may concern the likelihood of expression of a category, depending on type of referent, or type of construction.

Associative plurals are a special subtype of plural number which can be described as 'X and X's associates', e.g. Hungarian *Péter-ék* (Peter-ASSOCIATIVE.PLURAL) 'Peter and his family and/or friends and/or associates', Tariana *nami-sini* (father's.younger.brother-ASSOCIATIVE.PLURAL) 'father's younger brother and his associates'. ('Peter' and 'father's younger brother' are termed 'focal referents'.) In her seminal study of the phenomenon, Moravcsik (2003: 472) established the following implicational generalization which allows us to predict the marking of associative plural in a language:

(4) *The choice of the focal referent for associative plurals: a cross-linguistic generalization*
Proper name > Definite Kin noun > Definite Title noun > Other Definite Human noun
'If in a language, a nominal can be a focal referent of an associative plural, so can any other nominal to its left on the scale in that language.'

This generalization – stated as an implicational hierarchy – has predictive power. We have just seen that in Tariana an associative plural can be formed on a kin noun. Following the generalization in (4), it should also be possible to form associative plural on a proper name. This is indeed the case, e.g. Tariana *Marino-sini* 'Mario and his associates' (see also Moravcsik, Chapter 14 of this volume).

One of the best-known implicational hierarchies in linguistics is the Nominal Hierarchy, in Diagram 1.1.

The Nominal Hierarchy helps explain why in quite a few languages, nouns operate on an absolutive-ergative principle, and personal pronouns have nominative-accusative case-marking (phenomenon known as 'split ergativity': Dixon 1994, based upon Silverstein 1976). This is the case in Dyirbal, an Australian language. In (5), 'man' is the O marked

1st and 2nd person pronouns	Demonstratives 3rd person pronouns	Proper nouns	Common nouns	
			Human Animate	Inanimate

← more likely to be in A than O function

Diagram 1.1 Nominal Hierarchy (Dixon 1994: 85)

with the absolutive case, with zero realization. 'Woman' is the A (transitive subject) and is marked with the ergative case:

(5) [bayi yara]$_O$
NOMINAL.MARKER.ABS:MASC man:ABS
[ba-ŋgu-n yibi-ŋgu]$_A$ balga-n
NOMINAL.MARKER-ERG-FEM woman-ERGATIVE hit-PAST
'The woman hit the man'

In (6), 'woman' is the intransitive subject (S) and is marked with the absolutive case.

(6) [bala-n yibi]$_S$ miyanda-nyu
NOMINAL.MARKER.ABS-FEM woman:ABS laugh-PAST
'A woman laughed'

If A, S and O are expressed with pronouns, the marking changes. The same form of the pronoun 'I', *ŋaja*, is used for the S in the intransitive (7) and for A in transitive (8).

(7) ŋaja$_S$ bani-ñu
I:NOM come-PAST
'I came (here)'

(8) ŋaja$_A$ [bayi yara]$_O$ balga-n
I:NOM NOMINAL.MARKER.ABS:MASC man:ABS hit-PAST
'I hit a man'

In (9), a different form of the pronoun 'I' is used for an O. This is an accusative form:

(9) ŋayguna$_O$ [ba-ŋgu-n yibi-ŋgu]$_A$ balga-n
I:ACC NOMINAL.MARKER-ERG-FEM woman-ERG hit-PAST
'A woman hit me'

The principle behind split ergative marking based on the meaning of a noun phrase reflects a general principle: a participant in an unusual role may acquire special marking, and a participant in an expected role does not have to. First and second persons singular, 'I' and 'you', are the quintessential A (the 'agent', the 'perceiver', the 'donor' and so on). 'I' and 'you' are more likely to appear as A than as O. When they do appear in O function, they will be marked. Next most likely A are demonstratives and third person pronouns. Human, animate and inanimate nouns are less likely to be 'A'. Many verbs typically have a human noun in their A function (e.g. 'think', 'believe', 'tell'); for others, A tends to be human or animate ('bite', 'strike', 'see', 'hear'). There is more variety with regard to O. An inanimate noun can be O of most verbs.

The Nominal Hierarchy reflects our expectations as to which participants are more likely to appear in A than in O function. Participants at the left end of this hierarchy are more likely to be in A function and have

agentive properties. Those at the right end are more likely to be in O or S function; they can be conceived of as less agentive and less A-like. In Dixon's (1994: 85) words,

> it is plainly most natural and economical to 'mark' a participant when it is in an unaccustomed role. That is, we would expect that a case-marking language might provide morphological marking of an NP from the right-hand side of the hierarchy when it is in A function, and of an NP from the leftmost end when in O function (as an alternative to providing ergative marking for *all* A NPs, of whatever semantic type, or accusative marking for *all* O NPs).

The Nominal Hierarchy has explanatory power and also correlates with typical functions of nominals with different semantic features. It has also proved useful in other areas of language. In his pioneering work, Smith-Stark (1974) demonstrated its applicability to the expression of number: it is more likely that humans and animates will be overtly marked for number than inanimates (further work on this was done by Corbett 2000).

A well-known implication hierarchy in lexical typology concerns colour terminology in the world's languages. This was originally formulated by Berlin and Kay (1969: 5), as in Diagram 1.2.

This implicational hierarchy makes the following prediction: the presence of a term on the right implies the existence of all those to the left. So, a language with a basic term for 'blue' will be expected to have basic terms for green, yellow, red, white and black. The hierarchy is concerned with 'basic' colours (defined by Berlin and Kay 1969, and then refined in later studies).[13]

Implicational statements are formulated on the basis of language data. Once justified, they offer useful insights into the grammar of individual languages and provide an important investigative tool. A statistical analysis of a large corpus of Modern Russian showed that terms for 'black' and 'white' (the two foremost 'basic' colours) are the most frequent (Corbett 2011: 201–2). The three most basic colour terms in English – 'white', 'black' and 'red' – stand apart in that they can take the derivational suffix *en* as in *whiten, blacken* and *redden.*

This takes us to the next topic: the relations between linguistic typology and language analysis, in terms of its empirical base.

WHITE		GREEN			PURPLE
					PINK
	< RED		< BLUE	< BROWN	ORANGE
BLACK		YELLOW			GREY

Diagram 1.2 The colour terminology hierarchy originally suggested by Berlin and Kay (1969)

1.3 Linguistic Typology and Language Analysis

1.3.1 Inductive Nature of Typological Generalizations

Typological research expected to produce substantive and meaningful generalizations about languages has to be based on the analysis of languages themselves. That is, its foundations are bound to be inductive. This resonates with Bloomfield's statement (1933: 20):

> The only useful generalizations about language are inductive generalizations. Features which we think ought to be universal may be absent from the very next language that becomes accessible ... The fact that some features are widespread is worthy of notice and calls for an explanation; when we have adequate data about many languages, we shall have to return to the problem of general grammar and to explain these similarities and divergences, but this study, when it comes, will not be speculative but inductive.

Typology and language analysis feed into each other. In order to come up with sensible cross-linguistic generalization to be borne out by facts of languages, a typologist needs to rely on good-quality comprehensive reference grammars. The converse is also true. Detailed reference grammars of previously undescribed languages alert typologists to new phenomena and offer materials for new typological generalizations.

In a 1997 paper, DeLancey offered the first cross-linguistic snapshot of mirativity – a category whose main meaning is 'surprise' and 'unprepared mind'. Individual scholars of numerous languages had been aware of the existence of such a phenomenon. A similar range of mirative meanings (there termed 'admirative') occurs in a grammar of Albanian written in French by Dozon (1879: 226–7) (see Friedman 2003: 192–3, 213, and 2012 for the history, and the meanings of the term). A verb form with 'admirative' meaning was described in grammars of north-east Caucasian languages (e.g. Kibrik 1977, 1994). Many grammars of Quechuan languages also described a verb form with a major meaning of 'sudden realization or awareness' and 'surprise' (also known as 'sudden discovery tense': see Adelaar 1977, 2013; Adelaar with Muysken 2004).

Having a cross-linguistic overview of a mirative as more than one isolated exotic curiosity alerted grammarians across the world to its existence. The term 'mirative' and the notion 'mirativity' have gained currency over the past fifteen years, and new types of mirative systems (distinct from tense, aspect, modality and evidentiality) keep being discovered (see summaries in Aikhenvald 2012; DeLancey 2012).

Along similar lines, Nordlinger and Sadler's (2004) cross-linguistic study of nominal tense alerted grammarians to this category. More and more grammars turn their attention to it than previously. Having a special set of 'tenses' on nominals is no longer an exotic feature.

Kuteva's (1998) investigation of a poorly known cross-linguistic category 'action nearly averted' put the category 'on the map' (see also Overall, Chapter 15 of this volume). Thanks to these findings, our reference grammars are becoming fuller and more comprehensive – and our understanding of the limits of linguistic variation deepens.

The discipline of linguistic typology must rest upon firm empirical foundations. We now turn to the factual basis for typological hypotheses and generalizations.

1.3.2 The Basis for Typological Investigation

A typological study can vary in terms of its scope. It can range over a limited set of languages – all in one geographical region or all in one genetic family – or over all human languages. For typological generalizations and the typological classification to be significant, there should be a fair number of languages in the population considered – more than just a handful.

When a *typological* study aims to cover a large number of languages, potential problems arise. Perhaps as many as 4,000 languages are currently spoken or were recently spoken across the world. However, good and reliable reference grammars are available for only a fraction of these. The best way of getting as full a view of languages as possible is to try and access as many reliable sources as one can. The many partial and defective grammars are to be avoided in any typological research. And it is not the case that a newer source will necessarily be better, or more comprehensive than, an earlier one.[14]

Some typologists find it appropriate to limit themselves to an artificially constructed and limited sample. For instance, Bybee, Perkins and Pagliuca's (1994) survey of tense, aspect and modality in the languages of the world was based on a meagre sample of ninety-four languages from a limited set of genetic groupings, many of them not substantiated (such as 'Andean-Equatorial', 'Jê-Pano-Carib' and 'Indo-Pacific'). The grammars chosen, on a random basis, included numerous flawed and incomplete sources (see Dixon 2010a: 259–61, for further critique of sampling methodologies).[15]

It is useful to distinguish two approaches to typological study of grammatical categories and patterns. These can be referred to as Approach A and Approach B.

Approach A involves the following:

a. Working with primary sources. If some useful-looking information is discerned in a secondary source, this must always be checked back in the appropriate primary source. If no primary source is provided, the information in the secondary source cannot be taken account of.
b. Building typological generalizations on an inductive basis.

c. Searching out and using all reliable sources which relate to the topic under investigation. This will involve doing an initial survey of all languages of the world, and then homing in on those groups of languages where the topic under investigation is most manifested.
d. For every topic of investigation, taking account of social, areal and historical considerations, thus incorporating linguistic history into the patterns of distribution for the phenomenon. For instance, an open-ended typology of nasalization will pay attention to the fact that nasalization is a prominent areal feature of many parts of Amazonia but does not feature overmuch in European languages.
e. Attempting at an explanation of the nature and distribution of the phenomenon under investigation – language internal, historical and maybe also sociocultural.

It is typically the case that the proponents of Approach A have themselves undertaken intensive linguistic fieldwork and written grammars of previously undescribed languages. They thus have the experience which enables them to evaluate the worth of grammars.

Approach B typically involves the following:

a. Making copious use of secondary sources without attempting in all cases to check that they have quoted correctly from primary sources.
b. Sometimes working inductively, but sometimes deductively. The latter course involves putting forward an a priori hypothesis and then looking at just a few selected languages to see whether it can be upheld.
c. Consulting just a 'sample' of languages, without regard for whether they may be representative for the topic of investigation or whether the materials available are reliable.
d. Focusing on linguistic data as an end in itself, without paying too much attention to the social and areal contexts and historical origin.

Studies and surveys following Approach B – typically undertaken by those who have not undertaken original fieldwork themselves – run the risk of misinterpreting original sources. Those who aim at valid typological generalizations are better off avoiding such secondary investigations.

A typologist needs to exercise care in stating what is common and what is infrequent across languages. Suppletive classificatory verbs are a common feature in Athabaskan languages, and also in sign languages but uncommon in languages of most other areas. Haude (2010) presents nominal tense in Movima, an isolate from Bolivian Amazonia, as a typologically rare feature. This would be exotic if compared with Indo-European languages. However, Amazonian and Australian languages offer numerous examples of this phenomenon, enough to build a typology (see Nordlinger and Sadler 2004; Tonhauser 2007, 2008; summary in Aikhenvald 2012: 158–62). We will not be able to state the frequency of

any phenomenon with full assurance until most languages of the world have been properly investigated.

The idea has spread that using sampling techniques, and quantified correlations and chi-square tests and the like, makes linguistics 'scientific'. But what is needed, at the present time, is not quantitative comparison of superficial bits of surface structure, but rather qualitative study, providing careful and fine-grained analysis of the underlying structures of the multitude of languages which are still spoken and are in need of comprehensive documentation.

New sources shed new light on the distribution of particular categories and constructions cross-linguistically. Evidentiality (viewed as grammaticalized information source) was a rare bird in African languages. Recent comprehensive descriptions have revealed its existence in a number of groups (see König 2013 on Khoisan; Storch 2013 on Luwo; and a brief survey in Aikhenvald 2014).

What, then, is the optimal procedure for delimiting a database in typological enquiry? Is any type of sampling appropriate? The answer is: 'not really'. Instead, one should carry out an extensive survey, and then home in on intensive studies. For instance, when Alexandra Aikhenvald began working on a cross-linguistic study of evidentiality, her first step was to survey languages across every linguistic area and genetic family. The category is missing from certain regions, spottily represented in some, rampant in others. Having identified critical areas, she then embarked on a detailed examination of all the well-described languages there. She also looked at poorer grammars of some of the languages which might be expected to include an evidentiality system, seeing if she could spy any symptomatic features and, where possible, corresponded with authors of the grammars. Finally, as the result of six or seven years of fairly intensive study, in 2004 she published a 479-page monograph called *Evidentiality* which covered as much on the phenomenon as was available at the time.[16]

One vitally important point which has been mentioned before but deserves emphasis is that not everything which has appeared in print – been written down – is equally good. When embarking on a comparative study of the indigenous languages of the Americas, Peter S. Duponceau – in 1819 – set out his method of working:

> I left no book or manuscript unconsulted that came within my reach; but I examined the assertions of each writer with a critical eye, fully determined in no case to swear on the word of a master. I tried to discover the sources from which my authors had derived their knowledge; the opportunities which they had of acquiring it; the time which they had spent among the Indians, or in the study of their languages; the degree of attention which they had bestowed upon it, and the powers of mind by which they had been enabled to take a just and an accurate view of their subject. Finally, I rejected every thing that came in the shape of

> mere assertion, and paid attention only to those specimens of the different idioms in which their grammatical structure was sufficiently exhibited.

Such a comprehensive and critical approach will ensure the validity of typological generalizations. The time that will go into such work will be time well spent.

A word on the methodology and theoretical underpinnings of linguistic typology. Throughout the past two centuries, the bulk of inductive generalizations about the world's languages were cast in an analytic framework based on an in-depth description of linguistic facts not constrained by any ad hoc formal models. This framework has been recently given the name of 'basic linguistic theory' (see Dixon 2010a, 2010b, 2012). In this framework, every analytic decision has to be proved. It is oriented towards expanding our view of structural diversity. This is the perspective taken throughout this volume.[17]

1.3.3 Prerequisties for Consistent Analysis

The first and major prerequisite for a typological investigation is reliable sources. The quality of our typologies depends on the quality of the sources we have available, and it is indeed hard to compare a phenomenon in two languages or one which has a full comprehensive grammar with another that has just a small, barely informative sketch. This is a recurrent problem, highlighted by many authors of this volume.

Comprehensive reference grammars based on extensive fieldwork constitute an ideal solid basis for typological work (a list of examples of such grammars is in Dixon 2010a: 81–5). One needs to avoid partial descriptions based on limited sources for a language which is still well spoken. Partial and limited analyses based on field methods courses must be excluded from typological analysis.[18] A major commandment for a typologist – just as for any scholar – is to always go back to original sources (see Dixon 2010a: 64–6).

A further prerequisite for a coherent typological study is consistent and compatible analyses of languages compared. One can only sensibly compare things described in similar terms. Compatible analyses of languages are a sine qua non for typological comparison. A typologist examining grammars of a number of languages should not uncritically reproduce descriptions.

Different authors may use terms differently. Take the term 'serial verb construction'. This is generally applied to a sequence of verbs without any overt marking functioning as one predicate (see summary and references in Aikhenvald 2006). Scholars of Tupí-Guaraní languages (e.g. Jensen 1999) use this term to describe sequences of verbs, one of which is cast in a verb form marked for subordination (and termed 'gerund'). These structures,

no matter what they are called, should not be considered in a typology of serial verb constructions.

Descriptions of South-East Asian languages use the term 'classifier' to refer to a numeral classifier. In the Athabaskan linguistic traditions, 'classifier' is a grammatical voice marker on verbs. What Athabaskanists call 'gender' corresponds to verbal prefixes which classify a noun phrase in terms of its shape and form; this is very different from the way 'gender' is used in familiar Indo-European languages (see Aikhenvald, Chapter 12 of this volume).

The term 'passive' is used in a number of meanings. A survey of passives states: 'The analysis of the various constructions referred to in the literature as "passive" leads to the conclusion that there is not even one single property which all these constructions have in common' (Siewierska 1984: 1). This shows that a typologist must take great care in assessing what is described as a passive construction in each grammar under consideration. (For a comprehensive typology of passives and other valency reducing derivations, see Dixon 2010a: 166–8; 2012: 197–238).

When embarking on a typological investigation of category X, one must first of all adopt a working definition of 'what X is' and the types of criteria for recognizing it in a given language. Then, for each language which is studied, the available grammar(s) must be carefully assessed to see if what is there called 'X' accords to the definition and criteria (or, perhaps, whether something which is given a different label in the grammar does so). In some languages, what are called 'passive' will not satisfy the criteria and so should be excluded from the typological study.

A careful approach is particularly important in dealing with categories which are comparatively new in linguistic typology and may not be familiar to many general linguists. The category of 'frustrative' is a case in point. Simon Overall looks at the meanings of morphemes called 'frustrative' (and other morphemes, with similar overtones) (Chapter 15 of this volume) and provides a working definition which allows him to define the meanings and the uses of this hitherto marginalized category.

It would be a mistake to look up the index at the end of a grammar, see that, say, 'passives' or 'reciprocals' are discussed on pages 332–5 and then just read these four pages. As Antoine Meillet (1926: 16) put it, 'une langue constitue un système complexe de moyens d'expression, système où tout se tient' ('a language makes up a complex system of means of expression, a system in which everything holds together'). Scientific linguists who produce comprehensive grammars of languages naturally follow this tenet. Those who look at isolated bits of language, for some particular issue, go against this fundamental principle of systematic analysis. Analytic decisions about passives are likely to depend on or interrelate

with other kinds of analytic decisions, in various parts of the grammar (for example, concerning number and types of arguments). Proper procedure is to study the whole grammar in outline, to understand the context for this short discussion of passives, and then see how this links up with treatment of transitivity, other syntactic derivations which affect valency, discourse structure and the like.

In summary, one should undertake typological study of some aspect of the underlying organization of grammar, rather than of its surface realization solely guided by the use of a term. And one should carefully consider the analysis of this feature provided by each grammar that is included in the survey. In an ideal world, writers of grammars would all employ similar criteria and make analytic decisions on a similar set of criteria. But nothing is ideal.

A word on typology and reference grammars. Grammatical analysis and typological generalizations interact. A comprehensive reference grammar will be incomplete without a general typological profile of the language.

This highlights features that give the grammar its distinctive character, cast within a general perspective of cross-linguistic typological investigation. A typological profile needs to address parameters in morphological typology, mentioning – among many other topics:

- whether the language is prefixing, or suffixing, or both, and how many prefix and suffix positions are typically encountered; whether there are infixes;
- whether the language is isolating, or fusional, or agglutinating;
- whether the language is analytic or synthetic;
- whether the language is absolutive-ergative, or nominative-accusative (or combine both patterns);
- whether there are genders, or classifiers of any types, or both (more on this in Aikhenvald 2015: 30–2).

Typological profiles of languages are based on a set of synchronically attested features. However, in order to expand its explanatory powers, typology needs to go beyond a purely synchronic approach. We now turn to the relationships between typology, linguistic history and language contact.

1.4 Typology, History and Contact

Linguistic typology aims at establishing parameters for cross-linguistic variation at a particular point in time, without any consideration for language history, or contact between languages. That is, typology is synchronically oriented. Typological similarities have to be kept separate from shared features due to genetic inheritance, and changes induced by

language contact. Stating that languages are 'related' because they share a structural typological feature is nonsensical.[19]

However, typological research can branch out into historical linguistics and the investigation of diachronic language change. It can turn out to be crucial in setting limits for the potential of what one can expect in any language, including a putative proto-language.

What we learn through linguistic typology helps uncover what is likely, what is possible and what is unlikely in a human language. Linguistic typology sets a 'limit' to what a human language does. This limit, captured by the notion of 'typological plausibility', helps evaluate the likelihood of a linguistic reconstruction and the putative paths of linguistic change. We can thus work on a diachronic typology and a typology of historical processes going beyond the here-and-now of synchronic typology.

The 'typological plausibility principle' guides and limits what we should logically be able to reconstruct in our endeavour to recover the diachrony of a group of related languages. Typology can be used 'to justify the possibility of a given reconstruction by uncovering analogues in attested languages'. It can also provide 'an argument against a particular reconstruction by showing that there are either principles or strong empirical arguments that the reconstruction reflects a very unlikely language type, although the strength of such arguments will always depend on the level of development of typology' (Comrie 1993: 95). A further connection between typology and historical linguistics lies in establishing universals and universal tendencies in language change, including development of individual patterns (e.g. tones) and principles of grammaticalization (see also Hopper 1987; Hock 1991, 2010). Establishing limits to typological variation allows us to formulate predictions as to the ways in which a language might develop.

Similarities in the ways languages develop offer an additional link between historical investigation and typological research. Genetically related languages 'will pass through the same or strikingly similar phases': this 'parallelism in drift' (Sapir 1921: 171–2) accounts for additional similarities between related languages, even for those 'long disconnected'. Parallelism in drift may account for shared typological features of genetically related languages.

Typology as investigation of cross-linguistic variation is – ideally – independent of how languages in geographical proximity influence each other in contact situations. In practice, an informed study of cross-linguistic variation and its spread helps delineate linguistic areas and assign the weight of features (see Luraghi, Chapter 4 of this volume).

Each language community (save for a very few confined to a distant island or an inaccessible mountain valley) is in contact with other communities, speaking different dialects or languages. The communities will interact, through trade, shared festivals and rituals, inter-marriage and

maybe wars. Through all this, their languages also interact. They may come to sound more similar. They may borrow some vocabulary, and some structural and organizational features of the languages may converge, often resulting in a constellation of structural features which can be referred to as 'typological profile', or a 'type' of a linguistic area (see Enfield, Chapter 19, for languages of mainland Southeast Asia; Dixon, Chapter 20, for the Australian linguistic area; and Adelaar, Chapter 21, for Quechuan and Aymaran languages).

Thus, for instance, languages spoken in the Balkans share such features as (i) lack of the infinitive, (ii) syncretism of dative and genitive markers, (iii) postposed definite article and (iv) two-term evidentiality systems. This is subsumed under a term 'Balkan type'. A subfield of areal typology deals with linguistic features which spread and develop as a result of language contact. Areal typological studies are crucial for understanding the dynamics of possible language contact and language history (see, for instance, Aikhenvald and Dixon 1998, and further papers in Ramat 1998).

Typological profiles of linguistic families and linguistic areas constitute a fruitful avenue for examination, as they allow us to catch a glimpse of linguistic, and cognitive, diversity, and its limits. In addition, typological tendencies characteristic of a family can suggest the ways in which individual languages may change. Typological tendencies – and features – of an area may help us understand the dynamics of language contact, and contact-induced change.

The family or the area may be characterized by special features, such as possessive classifiers in Oceanic languages or multiple noun classes in Bantu languages. It is important to point out the parameters of variation within the family, with regard to any distinct typological feature. For instance, most Arawak languages of South America have two genders; however, Palikur has three genders, and Amuesha, Waurá and Terêna have lost gender distinctions altogether.

A set of synchronically established grammatical features found in a linguistic family or its subgroup can be referred to as its 'type'. Austronesian languages of the Philippines (and also some languages of Taiwan) have cross-linguistically unusual alternative transitive constructions whereby one argument is placed in 'focus'. This is shown by affix(es) on the verb indicating the function of the focused argument (which can be transitive subject, object or recipient, or an oblique (instrument or location). Having a system of such verbal 'focuses' or 'voices' is often referred to as the Philippine type.[20]

A feature or a property may be found in just one subgroup of a family. Take relational classifiers – a special class of morphemes used in constructions with alienable possession which characterize the possessed in terms of how it is handled (for instance, whether it can be eaten or drunk). By and large, these are found only in the Oceanic subgroup of

the Austronesian family (see Guérin, Chapter 29 of this volume). They can be referred to as an 'Oceanic-type' feature. Along similar lines, suppletive classificatory verbs are usually associated with Athabaskan languages where they are most prominent and best known. Classificatory postural verbs are a feature of the New Guinea Highlands (Aikhenvald, Chapter 12 of this volume).

Areal and historical studies help provide the answer to the question why the languages are the way they are. This is one of the goals of linguistic typology.

Now a word of caution. In order for a typology to be reliable, and relevant for historical and areal investigations, a scholar should be careful in establishing genetic affiliation of languages. Languages of the world are divided into proven families – such as Indo-European, Uralic, Dravidian, Tai-Kadai, Algonquian, Athabaskan, Arawak, Panoan, Carib, Tupí, to name just a few. Based on occasional shared formal similarities and shared typological features, speculative genetic groupings have been suggested (such as Indo-Pacific, Ge-Pano-Carib, Macro-Equatorial, Amerind, Arawakan, Pama-Nyungan, Nostratic, etc.). Putative long-distance groupings – unless proved – remain, in Matthews' (2007: 268) words, 'the kind of hypothesis one believes to the extent that one believes in that kind of hypothesis'. To make a typology a scientific enterprise, such science-fictional units need to be avoided.[21]

1.5 About This Volume

This volume offers a state-of-the art view of major issues within present-day linguistic typology in its varied guises, with equal attention to (a) universal tendencies across languages in various domains, (b) typological variation within individual categories and (c) patterns of linguistic diversity, across a selection of language families and groups from across the world.

The volume consists of three parts. Part I, 'Domains of Linguistic Typology', comprises nine chapters. The first five chapters cover a subfield of linguistic typology and its interaction with other fields or areas within the discipline of linguistics: Harry van der Hulst (Chapter 2) offers a reappraisal of typological research in phonology. Thomas E. Payne (Chapter 3) addresses various issues in morphological typology, including its history and development over the past centuries. Silvia Luraghi (Chapter 4) discusses the relationship between typology, historical and comparative linguistics, and language contact and change. Grammaticalization is a major historical and synchronic mechanism of creation and development of grammatical forms. Sociolinguistic typology is a relatively recent arrival on the linguistic scene. This is concerned with establishing types of social structures and the ways in which

these are reflected in linguistic practices, including language attitudes, language use and sociolinguistically based variation. In Chapter 5, Peter Trudgill – in many ways the founder of the discipline – offers a discussion of sociolinguistic typology with regard to interdependencies between social structures and linguistic complexity, demonstrating the explanatory power of such correlations. In Chapter 6, Heiko Narrog discusses the ways in which tendencies and generalizations in linguistic typology interact with principles of grammaticalization.

The last four chapters in Part I are concerned with typological properties of a selection of different linguistic systems rarely addressed in the typological literature. Sign languages occupy a different modality, with different means of expression from spoken languages. Only by investigating both spoken and signed languages can we come to understand the full range of possibilities of human language structures. Typological studies of sign languages is a newly emerging field which is likely to open new paths into cross-linguistic studies and linguistic diversity beyond the spoken word. Ulrike Zeshan can be rightfully considered the pioneer of the field (e.g. Zeshan 2004). Much of modern-day typology has focused on the properties of spoken languages. To amend this historical bias, Ulrike Zeshan and Nick Palfreyman (Chapter 7) offer an incisive typological sketch of sign languages and some of the differences and commonalities with spoken languages.

Mixed, or intertwined languages – a number of which arose as a result of conscious language manipulation – are a fruitful and yet largely unexplored field of typological research. Typological features of mixed languages are the topic of Chapter 8, by Peter Bakker. In terms of their origins and functions, Creole languages are a special kind: they arose as a consequence of people of several different linguistic backgrounds having to communicate with each other and developing a common, often simplified, linguistic variety. Creole languages have a distinct set of structural features, and it is a fascinating task for a typologist to identify them and correlate them with the donor languages and processes involved. Typological features of Creole languages are the topic of Chapter 9, by Aymeric Daval-Markussen and Peter Bakker.

What are the limits of human linguistic creativity? Secret languages – used for ritual communication or to exclude an unwanted group – and linguistic taboos involve manipulating linguistic structures in comparable ways; the typological principles behind these and their place within language structure are addressed by Anne Storch, in Chapter 10.

Part II, 'Typology of Grammatical Categories', comprises eight contributions, each dealing with typological investigation of one core aspect of grammatical structure. In Chapter 11, David Beck offers

a comprehensive typology of morphological means (in some ways complementing Chapter 3, on morphological typology). In Chapter 12, Alexandra Y. Aikhenvald offers a typology of noun categorization devices (which cover genders, noun classes and classifiers of different kinds). The typology of negation – a universal category found in every human language – is the topic of Matti Miestamo's Chapter 13. Edith A. Moravcsik offers a comprehensive reappraisal of expression and meanings of number and quantity across the world's languages in Chapter 14.

Some categories within Part II are relatively recent arrivals on the linguistic scene. The verbal category of frustrative is a prominent feature of many languages of Lowland Amazonia. It expresses non-realization of an expected outcome – and also the frustration of the speaker. In Chapter 15, Simon E. Overall offers a typology of frustrative marking, its semantics and development in Amazonian languages, as a first step towards a comprehensive outline of this fascinating, and so far neglected, category.

A number of African languages in what is known as the 'Macro-Sudanic belt' have a paradigm of person markers indicating co-reference with the real or imagined author or source of secondary discourse. This has come to be known as 'logophoricity', a type of marking rarely found outside West Africa. In Chapter 16, Felix K. Ameka offers a comprehensive typological analysis of logophoricity, its meanings and expression, with special attention to the sociocultural and communication practices in West Africa which underlie logophoricity.

Switch reference – a major mechanism for keeping track of participants in discourse – is a salient feature of many languages of the world, especially in New Guinea and North and South America. Chapter 17, by John Roberts, contains a concise typological appreciation of cross-linguistically common patterns of switch reference (with special focus on the Papuan languages of New Guinea, the major expertise of the author). Chapter 18, by Eric Pederson, is somewhat different in nature and scope: it contains an up-to-date overview of approaches to the expression and analysis of the meaning of one semantic category – motion events – across human languages.

We place special focus on the analysis of the world's languages within a typological perspective, thus contributing to the appreciation and providing an up-to-date statement of linguistic diversity and variation between languages. Part III, 'Typological Profiles of Language Families and Linguistic Areas', features twelve contributions, each dedicated to a typological profile of a selected linguistic area or a language family, or a subgroup. We have striven to achieve a combination of well-known and not-so-well-known areas and families, to give our readers a glimpse into the range of typological diversity across the world's languages. Each typological profile of a language family or a subgroup is synchronically

oriented, with a mention of features reconstructible to the proto-language.

The first three chapters address typological features of linguistic areas. In Chapter 19, N. J. Enfield presents an almost exhaustive analysis of the features of languages in the mainland Southeast Asia area. The Australian linguistic area and its linguistic traits are discussed by R. M. W. Dixon in Chapter 20. In Chapter 21, Willem F. H. Adelaar provides a comprehensive account of the major typological features of Aymaran and Quechuan languages in the Central Andean region of South America.

The remaining chapters focus on typological properties of a selection of genetically established families and subgroups. We start with North America and proceed from north to south. In Chapter 22, Michael Fortescue offers a comprehensive account of the Eskimo-Aleut language family. Keren Rice and Willem de Reuse offer a detailed, in-depth account of the Athabaskan (Dene) language family in Chapter 23.[22] An all-embracing account of the Iroquoian language family and its daunting complexities is presented in Chapter 24, by Marianne Mithun. It is hard to give justice to the linguistic diversity of Amazonia. Facts from Amazonian languages are discussed in many of the chapters in Parts I and II. Part III contains just one contribution dealing with an Amazonian group: in Chapter 25, Elena Mihas offers an in-depth account of the typological characteristics of the Kampa subgroup of Arawak languages (one of the best-established divisions within this large family, some of whose internal classification remains a matter for investigation). The Omotic language family, described in detail by Azeb Amha in Chapter 26, is part of a large Afroasiatic grouping. The genetic unity, and the status of Omotic languages as a family within Afroasiatic, was recognized later than that of its other members. The Semitic language family – the family with the longest history of attestation within Afroasiatic which boasts some of the world's earliest documented languages such as Hebrew and Akkadian – is addressed in Chapter 27, by Aaron D. Rubin. The many typologically interesting features of the Dravidian language family are the topic of Chapter 28, by Sanford Steever. Austronesian languages comprise one of the largest families of the world. Its Oceanic subgroup – consisting of about 500 languages – has numerous distinctive features examined by Valérie Guérin, in Chapter 29. The island of New Guinea is a high spot of linguistic diversity in every possible meaning of the term. Papuan languages have been discussed throughout a number of chapters in Parts I and II. Chapter 30, by Lourens de Vries, gives an in-depth account of the relatively lesser known Greater Awyu Ndumut language family of West Papua, their history and distinctive properties.

We have assembled a stellar team of authors experienced in linguistic typology, historical comparative issues and first-hand investigation of the relevant languages. Throughout the volume, the analysis is cast in

terms of basic linguistic theory – the cumulative typological functional framework in terms of which almost all reference grammars are cast, and empirically sound generalizations – which will stand the test of time – formulated.

It would be nonsensical to think of this – or any other – volume on linguistic typology as 'the last word'. We hope that this volume will contribute to our understanding of what is common to many languages, and how human languages vary, expanding our knowledge and appreciation of the wondrous world of linguistic diversity and the surprises it has in stock.

Notes

1. Further discussion of the subject matter of linguistic typology can be found in Comrie (1988, 1989, 1991), Bazell (1958), Mallinson and Blake (1981), Daniel (2011) and Greenberg (1974), and references there.
2. See Evans and Levinson (2009), Comrie (1989) and references there on the limits of 'absolute' universals. Predictive powers of implicational universals are discussed by Corbett (2011) and especially Moravcsik (2011). Implicational universal tendencies can be stated in the form of hierarchies; more on this in §1.2.3. See Ohala (1975) for further discussion of nasal vowels and nasality.
3. See the discussion in Aikhenvald (2000: 244–5); see Plank and Schellinger (1997) for further counterexamples. Comrie (1989: 17–23) outlines a distinction between formal and substantive universals; see an overview in Cristofaro (2011). Moravcsik (2011) offers a useful attempt at explaining linguistic universals.
4. See Aikhenvald (2006) on serial verb constructions, Smith-Stark (1974) on number marking.
5. Approaches to the notion of universals vary depending on linguists' theoretical persuasion. One extreme – which appears in Berman and Slobin (1994: 641) – is to view 'each individual language' as 'representing but one variant of a familiar and universally human pattern'. The other extreme is to foreground diversity between languages at the expense of potential universal or near-universal hypotheses (for example, by Evans and Levinson 2009, reflected in their title 'the myth of linguistic universals'). By working together with experts on individual languages and especially fieldworkers, linguistic typologists are making progress towards understanding and explaining the limits of variation between languages.
6. The term 'parameter' is used here to refer to a general characteristic of languages. Note that this term has been used in a somewhat different sense in generative formalist linguistic literature, where a parameter is understood as a specific property predicting a number of other specific

properties. A general critique of the theory of parameters can be found in Newmeyer (1998: 350–64).

7. See Payne, Chapter 3 of this volume, for a comprehensive analysis of advances in morphological typology.
8. See Dixon (2010a: 325–8), Aikhenvald (2007) on morphological types of languages; Fortescue (forthcoming) on the notion of polysynthesis.
9. The foundations of linguistic typology as a principled study of cross-linguistic differences and similarities, and as linguistic classification based on explicit parameters were laid in the early nineteenth century. Early typologies were oriented towards classifying languages in terms of morphological complexity of words and their composition. The best-known early typology is that by Wilhelm von Humboldt (1836), based on earlier work by August Wilhelm von Schlegel (1818) and Friedrich von Schlegel (1808). An evolutionary perspective was added by Schleicher (1859). Sapir's work can be considered a further breakthrough in cross-linguistic study of languages and their classification based on structural properties (see Sapir 1921). See historical surveys in Comrie (1981), Greenberg (1974), Ramat (2011), Haase (2001a, 2001b), Morpurgo Davies (1975, 1997) and Kemmer (2003).
10. Since morphological typology and constituent order typology address the language as a whole, they are sometimes referred to as 'holistic', in contrast to typologies of structures, categories or construction types which can be called 'partial' (cf. Comrie 1991, 2001). Typology of constituent order (mislabelled as 'word order typology') has dominated the field of typology, with variable success, since the 1960s. Further discussion can be found in Comrie (1989, 1991), Mallinson and Blake (1981), Moravcsik (2011) and Dryer (2007); see criticism in Vennemann (1974) and Hawkins (1983, 2001).
11. See also Comrie (1991: 444); Hewitt (1979: 36) for Abkhaz (glosses partially from Comrie 1991: 444).
12. See also Corbett (2011), Comrie (1989), Moravcsik (2003) on associative plural; Smith-Stark (1974) and Dixon (1994) for the Nominal Hierarchy.
13. Further refinements can be found in Kay and McDaniel (1978), Kay, Berlin and Merrifield (1991) and Kay et al. (2009); see also Uusküla and Sutrop (2010) for discussion of a possible exception. Foley (1998: 150–65) contains a useful summary. Further well-known hierarchies include the agreement hierarchy (Corbett 1979), which concerns agreement on the predicate, attribute, relative and personal pronoun; and the Accessibility Hierarchy which addresses the noun phrase positions which can be relativised (Keenan and Comrie 1977). Not every hierarchy or correlation holds out when addressed in some detail. For instance, Greenberg's (1972) general hypothesis that 'if a language has numeral classifiers it does not have obligatory number on nouns' has been proved erroneous (it holds only for languages of

highly analytic and isolating profile) (see Aikhenvald 2000: 100–1; Nomoto 2013).

14. Anyone wanting to obtain information on the Australian language Gumbaynggirr should consult not only Eades (1979) but also a fuller account in Smythe (1948/9). Along similar lines, Morse and Maxwell's (1999) grammar of Kubeo, a Tucanoan language, is perhaps not ideal, but it covers many more topics in much more depth than the more recent one, by Chacon (2012).
15. A further ambitious, and yet disappointing, attempt at offering a typological picture of a variety of grammatical and other categories across the world is *The world atlas of language structures* (WALS) (Haspelmath et al. 2005). This source suffers from errors in quoting and interpreting data sources, variable coverage for different parameters and idiosyncratic criteria (see criticism in Schultze 2007; Bright 2007; Plank 2009 (and other papers in the same issue); Dixon 2012: 87, 290, 338, 462–3; Dixon 2010b: 258).
16. As new systems and descriptions become available, this work goes on; major outcomes are summarized in Aikhenvald, *The Oxford handbook of evidentiality* (forthcoming).
17. Deduction-based investigations, such as the generative theory of universal grammar, have been of limited use to our understanding of cross-linguistic variation of languages, and their potential; see Kemmer (2003: 312) and Haider (2001) on the limited contribution to typological research made by formal theories. These approaches are usually based on data from relatively few languages, focusing on grammaticality judgements (rather than language use); they are often oriented towards English and tend to seek explanation for the hypotheses in theory-specific structural properties which are assumed to be common to all languages.
18. A typologist ought to be particularly wary of studies based on translations, limited corpora and especially field methods courses. Kara is an Oceanic language spoken by a few thousand people in New Ireland. There is no comprehensive grammar of Kara as yet. Dryer (2013) is a 'grammar' of Kara based on the analysis of the translation of the New Testament into the language. This source can perhaps be used for translation studies, but should not be used for typological comparison (because it does not reflect the language as it is spoken or used by the community). Similarly, Bowern's (2011) description of Titan, a major indigenous language of the Manus province of Papua New Guinea, is based on the analysis of texts collected by missionaries. This is not an appropriate source for any comparative studies, because of its limited scope. Loughnane's (2003, 2005) study of Golin is based on a field methods class and cannot be used in typological research (a scholar interested in related languages will be better off perusing a comprehensive grammar of Dom, by Tida 2006). Similarly, Lawler's

(1988) study of passives in Acenhese is compromised by the limited data obtained during a field methods course. Interested scholars ought to use Durie's (1985, 1988) work.

19. See, for instance, Greenberg (1974), Dixon (1997), Hock (1991), among other sources, on the independence of typological and genetic classifications of languages.
20. See Blust (2008) for a Pan-Austronesian perspective on 'Philippine-type' focus or 'voice' focus; and further references and analysis in Dixon and Aikhenvald (1997).
21. Typologists – or any linguistic scholars – worth their salt will also base their statements on genetic affiliation of languages on the opinions of respected specialists in the field – that is, Indo-Europeanists on Indo-European, Uralicists on Uralic and so forth. *Glottolog* (www.glottolog.org, Hammarström et al. 2015) is a newly established resource which is highly problematic. The first-named co-author of this chapter whose major expertise and publishing record lies in Arawak (also known as Maipuran) languages and other languages of South America was overwhelmed with the sheer number of mistakes and misinterpretations in the fanciful classification of languages of this family, indiscriminate use of linguistic references of mixed quality. There is, in addition, a general disregard for specialist scholarly achievements in the field and the work of experts. Numerous mistakes and misinterpretations have been found by the second co-author, in sections on Australian languages. Until such time as these issues are addressed, Glottolog will remain a source to be used with extreme caution. A good general source on the world's languages is Brown and Ogilvie (2009).
22. The name of the language family under discussion is spelled in several ways: Athabaskan, Athapaskan, and Athabascan. Recently, the term 'Dene' (the word for 'person' in many of the languages) has been used for the name of the family.

References

Adelaar, Willem F. H. 1977. *Tarma Quechua: Grammar, texts, dictionary*. Lisse: The Peter de Ridder Press.

2013. A Quechuan mirative? In Alexandra Y. Aikhenvald and Anne Storch (eds.), *Perception and cognition in language and culture*, pp. 95–110. Leiden: Brill.

Adelaar, Willem F. H. with Pieter Muysken. 2004. *The languages of the Andes*. Cambridge University Press.

Aikhenvald, Alexandra Y. 2000. *Classifiers: A typology of noun categorization devices*. Oxford University Press.

2004. *Evidentiality*. Oxford University Press.

2006. Serial verb constructions in a typological perspective. In Alexandra Y. Aikhenvald and R. M. W. Dixon (eds.), *Serial verb constructions: A cross-linguistic typology*, pp. 1–87. Oxford University Press.
2007. Typological dimensions in word formation. In Timothy Shopen (ed.), *Language typology and syntactic description*, Vol. III, pp. 1–65. Cambridge University Press.
2012. The essence of mirativity. *Linguistic Typology* 16(3): 435–85.
2014. The grammar of knowledge in typological perspective. In Alexandra Y. Aikhenvald and R. M. W. Dixon (eds.), *The grammar of knowledge: A cross-linguistic typology*, pp. 1–51. Oxford University Press.
2015. *The art of grammar*. Oxford University Press.
(ed.). Forthcoming. *The Oxford handbook of evidentiality*. Oxford University Press.
Aikhenvald, Alexandra Y. and R. M. W. Dixon. 1998. Evidentials and areal typology: A case study from Amazonia. *Language Sciences: Areal typology* (special issue, ed. Paulo Ramat) 20: 241–57.
Bazell, C. E. 1958. Linguistic typology. An inaugural lecture delivered on 26 February 1958. School of Oriental and African Studies, University of London.
Berlin, Brent and Paul Kay. 1969. *Basic color terms: Their universality and evolution*. Los Angeles: University of California Press.
Berman, Ruth A. and Dan I. Slobin. 1994. *Relating events in narrative: A crosslinguistic developmental study*. Hillsdale, NJ: Lawrence Erlbaum.
Bloomfield, Leonard. 1933. *Language*. New York: Holt.
Blust, Robert. 2008. *Austronesian languages*. Canberra: Pacific Linguistics.
Bowern, Claire. 2011. *Sivisa Titan: Sketch grammar, texts, vacabulary based on material collected by P. Josef Meier and Po Minis*. Honolulu: University of Hawai'i Press.
Bright, William. 2007. Review of *The world atlas of language structures*, edited by Martin Haspelmath, Matthew S. Dryer, David Gil and Bernard Comrie. *International Journal of American Linguistics* 73: 241–54.
Bybee, Joan L., Revere D. Perkins and William Pagliuca. 1994. *The evolution of grammar: Tense, aspect, and modality in the languages of the world*. University of Chicago Press.
Brown, Keith and Sarah Ogilvie (eds.). 2009. *Concise encyclopedia of languages of the world*. Oxford: Elsevier.
Chacon, Thiago. 2012. The phonology and morphology of Kubeo: The documentation, theory, and description of an Amazonian language. PhD dissertation, University of Hawai'i, Manoa.
Comrie, Bernard. 1981. *Language universals and linguistic typology*. University of Chicago Press.
1988. Linguistic typology. *Annual Review of Anthropology* 17: 145–59.

1989. *Language universals and linguistic typology: Syntax and morphology*, 2nd revised edn. Oxford: Blackwell and University of Chicago Press.
1990. Holistic versus partial typologies. In Werner Bahner, Joachim Schildt and Dieter Viehweger (eds.), *Proceedings of the Fourteenth International Congress of Linguists*, Vol. I, pp. 139–48. Berlin: Akademie-Verlag.
1993. Typology and reconstruction. In Charles Jones (ed.), *Historical linguistics: Problems and perspectives*, pp. 74–97. London: Longman.
2001. Different views of language typology. In Martin Haspelmath, Ekkehard König, Wulf Oesterreicher and Wolfgang Raible (eds.), *Language typology and language universals: An international handbook*, Vol. I, pp. 24–39. Berlin: Walter de Gruyter.
Corbett, Greville G. 1979. The agreement hierarchy. *Journal of Linguistics* 15: 203–24.
2000. *Number*. Cambridge University Press.
2011. Implicational hierarchies. In Jae Jung Song (ed.), *The Oxford handbook of linguistic typology*, pp. 190–205. Oxford University Press.
Cristofaro, Sonia. 2011. Language universals and linguistic knowledge. In Jae Jung Song (ed.), *The Oxford handbook of linguistic typology*, pp. 227–79. Oxford University Press.
Daniel, Michael. 2011. Linguistic typology and the study of language. In Jae Jung Song (ed.), *The Oxford handbook of linguistic typology*, pp. 43–68. Oxford University Press.
DeLancey, Scott. 1997. Mirativity: The grammatical marking of unexpected information. *Linguistic Typology* 1: 33–52.
2012. Still mirative after all these years. *Linguistic Typology* 16: 529–64.
Dixon, R. M. W. 1982. *Where have all the adjectives gone? And other essays in semantics and syntax*. Berlin: Mouton.
1994. *Ergativity*. Cambridge University Press.
1997. *The rise and fall of languages*. Cambridge University Press.
2004. Adjective classes in typological perspective. In R. M. W. Dixon and Alexandra Y. Aikhenvald (eds.), *Adjective classes: A cross-linguistic typology*, pp. 1–49. Oxford University Press.
2010a. *Basic linguistic theory*, Vol. I: *Methodology*. Oxford University Press.
2010b. *Basic linguistic theory*, Vol. II: *Grammatical topics*. Oxford University Press.
2012. *Basic linguistic theory*, Vol. III: *Further grammatical topics*. Oxford University Press.
Dixon, R. M. W. and Alexandra Y. Aikhenvald. 1997. A typology of argument-determined constructions. In Joan L. Bybee, John Haiman and Sandra Thompson (eds.), *Essays on language function and language type*, pp. 71–113. Amsterdam: John Benjamins.
Dozon, A. 1879. *Manuel de la langue chkipe ou albanaise*. Paris: Ernest Leroux.
Dryer, Matthew. 2007. Word order. In Timothy Shopen (ed.), *Language typology and syntactic description*, Vol. I: *Clause structure*, pp. 61–131. Cambridge University Press.

2013. *A grammatical description of Kara-Lemakot*. Canberra: Asia-Pacific Linguistics.

Duponceau, Peter S. 1819. Report of the Corresponding Secretary to the committee of his progress in the investigation committed to him of the General Character and Forms of the Languages of the American Indians – Read, 12th Jan, 1819. In *Transactions of the Historical and Literary Committee of the American Philosophical Society, held at Philadelphia, for promoting useful knowledge*, Vol. I, pp. xvii–xlvi. Philadelphia: Abraham Small.

Durie, Mark. 1985. *A grammar of Acehnese: On the basis of a dialect of North Aceh*. (Verhandelingen van het Koninklijk Instituut voor Taal-, Land- en Volkenkunde). Amsterdam: Foris Publications.

1988. The so-called passive of Acehnese. *Language* 61: 104–13.

Eades, Diana. 1979. Gumbaynggirr. In R. M. W. Dixon and Barry J. Blake (eds.), *Handbook of Australian languages*, Vol. I, pp. 245–361. Canberra: Australian National University Press; Amsterdam: John Benjamins.

Evans, Nicholas and Stephen C. Levinson. 2009. The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences* 32: 429–92.

Foley, William A. 1998. *Anthropological linguistics: An introduction*. London: Blackwell Publishers.

Fortescue, Michael. Forthcoming. Polysynthesis. In Nicholas Evans, Michael Fortescue and Marianne Mithun (eds.), *The Oxford handbook of polysynthesis*. Oxford University Press.

Friedman, Victor A. 2003. Evidentiality in the Balkans with special attention to Macedonian and Albanian. In Alexandra Y. Aikhenvald and R. M. W. Dixon (eds.), *Studies in evidentiality*, pp. 189–218. Amsterdam: John Benjamins.

2012. Perhaps mirativity is phlogiston, but admirativity is perfect: On Balkan evidential strategies. *Linguistic Typology* 16: 505–27.

Greenberg, Joseph H. 1963. Some universals of grammar with particular reference to the order of meaningful elements. In Joseph H. Greenberg (ed.), *Universals of language*, pp. 58–90. Cambridge, MA: MIT Press.

1966. *Language Universals, with special reference to feature hierarchies*. (Janua linguarum, Series Minor, 59). The Hague: Mouton.

1972. Numeral classifiers and substantival number: Problems in the genesis type. *Working Papers in Language Universals*; repr. in Greenberg 1990 (pp. 16–93).

1974. *Language typology: A historical and analytic overview*. The Hague and Paris: Mouton.

1990. *On language: Selected writings of Joseph H. Greenberg*, ed. K. Denning and S. Kemmer. Stanford University Press.

Haase, Martin. 2001a. *Sprachtypologie bei Edward Sapir*. In Martin Haspelmath, Ekkehard König, Wulf Oesterreicher and Wolfgang Raible (eds.),

Language typology and language universals: An international handbook, Vol. I, pp. 264–6. Berlin: Walter de Gruyter.

2001b. Sprachtypologie und Universalienforschung bei Joseph H. Greenberg. In Martin Haspelmath, Ekkehard König, Wulf Oesterreicher and Wolfgang Raible (eds.), *Language typology and language universals: An international handbook*, Vol. I, pp. 280–3. Berlin: Walter de Gruyter.

Haider, Hubert. 2001. Parametrisierung in der generativen Grammatik. In Martin Haspelmath, Ekkehard König, Wulf Oesterreicher and Wolfgang Raible (eds.), *Language typology and language universals: An international handbook*, Vol. I, pp. 283–93. Berlin: Walter de Gruyter.

Hammarström, Harald, Robert Forkel, Martin Haspelmath and Sebastian Bank (eds.). 2015. *Glottolog 2.5*. Leipzig: Max Planck Institute for Evolutionary Anthropology.

Haspelmath, Martin, Matthew S. Dryer, David Gil and Bernard Comrie (eds.). 2005. *The world atlas of language structures*. Oxford University Press.

Haude, Katarina. 2010. 'She kisses her late husband' = 'She kissed her husband': nominal tense in Movima. In Jan Wohlgemuth and Michael Cysouw (eds.), *Rara and rarissima: Documenting the fringes of linguistic diversity*, pp. 95–116. Berlin: De Gruyter Mouton.

Hawkins, John A. 1983. *Word order universals*. New York: Academic Press.

2001. The role of processing principles in explaining language universals. In Martin Haspelmath, Ekkehard König, Wulf Oesterreicher and Wolfgang Raible (eds.), *Language typology and language universals: An international handbook*, Vol. I, pp. 360–9. Berlin: Walter de Gruyter.

Heine, Bernd and Tania Kuteva. 2001. *World lexicon of grammaticalization*. Cambridge University Press.

Hewitt, B. G. 1979. *Abkhaz*. (Lingua Descriptive Studies 2). Amsterdam: North Holland Publishing Company.

Hock, Hans H. 1991. *Principles of historical linguistics*, 2nd edn. Berlin: Mouton de Gruyter.

2010. Typology and universals. In Silvia Luraghi and Vit Bubenik (eds.), *Bloomsbury companion to historical linguistics*, pp. 59–69. London: Bloomsbury.

Hopper, Paul J. 1987. Principles of grammaticization: Towards a diachronic typology. In W. P. Lehmann (ed.), *Language typology 1987: Systematic balance in language – Papers from the Linguistic Typology Symposium, Berkeley, 1–3 December 1987*, pp. 167–80. Amsterdam: John Benjamins.

Humboldt, Wilhelm von. 1836. *Über die Verschiedenheit des menschlichen Sprachbaues und ihren Einfluss auf die geistige Entwickelung des Menschengeschlechts*. Berlin: Druckerei der Königlichen Akademie der Wissenschaften.

Jensen, Cheryl. 1999. Tupí-Guaraní. In R. M. W. Dixon and Alexandra Y. Aikhenvald (eds.), *The Amazonian languages*, pp. 125–64. Cambridge University Press.

Kay, Paul and Chad K. McDaniel. 1978. The linguistic significance of the meanings of basic color terms. *Language* 54: 610–46.

Kay, Paul, Brent Berlin and William R. Merrifield. 1991. Biocultural implications of color naming. *Journal of Linguistic Anthopology* 1: 12–25.

Kay, Paul, Brent Berlin, Luisa Maffi, William R. Merrifield and Richard Cook, 2009. *World color survey*. Stanford: Center for the Study of Language and Information.

Keenan, Edward L. and Bernard Comrie. 1977. Noun phrase accessibility and universal grammar. *Linguistic Inquiry* 8: 63–99.

Kemmer, Suzanne. 2003. Typology and universals. In William J. Frawley (ed.), *International Encyclopedia of Linguistics*, 2nd edn, pp. 310–15. Oxford University Press.

Kibrik, A. E. 1977. *Opyt strukturnogo opisanija archinskogo jazuka. Tom II. Taksonomicheskaja grammatika* [An essay in structural description of Archi. Volume II. Taksonomic grammar]. Moscow: Izdateljstvo Moskovskogo Universiteta.

1994. Archi. In Rieks Smeets (ed.), *The indigenous languages of the Caucasus*, Part 2, pp. 297–365. New York: Caravan Books.

König, Christa. 2013. Source of information and unexpected information in !Xun: evidential, mirative and counterexpectation markers. In Alexandra Y. Aikhenvald and Anne Storch (eds.), *Perception and cognition in language and culture*, pp. 65–84. Leiden: Brill.

Kuteva, Tania. 1998. On identifying an evasive gram: Action narrowly averted. *Studies in Language* 22: 113–60.

Lawler, John M. 1988. On the questions of Acenhese 'passive'. *Language* 64: 114–17.

Loughnane, Robyn. 2003. Reported speech in Golin (a Papuan language of New Guinea). BA Honours thesis, University of Melbourne.

2005. Reported speech constructions in Golin. In Nicholas Evans, Jutta Besold, Hywel Stoakes and Alan Lee (eds.), *Materials on Golin: Grammar, texts and dictionary*, pp. 132–51. Department of Linguistics and Applied Linguistics, University of Melbourne.

Mallinson, Graham and Barry J. Blake. 1981. *Language typology: Cross-linguistic studies in syntax*. Amsterdam: North Holland Publishing Company.

Matthews, P. H. 2007. *Concise dictionary of linguistics*. Oxford University Press.

Meillet, Antoine. 1926. *Linguistique historique et linguistique génerale*. Paris: Champion.

Moravcsik, Edith. 2003. A semantic analysis of associative plurals. *Studies in Language* 27: 469–503.

2011. Explaining linguistic universals. In Jae Jung Song (ed.), *The Oxford handbook of linguistic typology*, pp. 69–89. Oxford University Press.

Morpurgo Davies, A. 1975. Language classification in the nineteenth century. In T. A. Sebeok (ed.), *Current trends in linguistics*, Vol. 13: *Historiography of linguistics*, pp. 606–716. The Hague: Mouton.

1997. *The history of linguistics*, Vol. IV: *Nineteenth-century linguistics*. London: Longman.

Morse, Nancy L. and Michael B. Maxwell. 1999. *Cubeo grammar. Studies in the languages of Colombia 5*. Arlington: Summer Institute of Linguistics; University of Texas at Arlington.

Newmeyer, Frederick J. 1998. *Language form and language function*. Cambridge, MA: MIT Press.

Nichols, Johanna. 1986. Head-marking and dependent-marking grammar. *Language* 62: 56–119.

Nomoto, Hiroki. 2013. Number in classifier languages. PhD dissertation, University of Minnesota.

Nordlinger, Rachel and Louisa Sadler. 2004. Nominal tense in cross-linguistic perspective. *Language* 80: 776–806.

Ohala, John J. 1975. Phonetic explanations for nasal sound patterns. In C. A. Ferguson, L. M. Hyman and J. J. Ohala (eds.), *Nasalfest: Papers from a symposium on nasals and nasalization*, pp. 289–316. Stanford: Language Universals Project.

Plank, Frans. 2009. WALS values evaluated. *Linguistic Typology* 13: 51–76.

Plank, Frans and Wolfgang Schellinger. 1997. The uneven distribution of genders over numbers: Greenberg Nos. 37 and 45. *Linguistic Typology* 1: 53–101.

Ramat, Paulo (ed.). 1998. *Areal typology*. Special issue of *Language Sciences* 20(3).

2011. The (early) history of linguistic typology. In Jae Jung Song (ed.), *The Oxford handbook of linguistic typology*, pp. 9–24. Oxford University Press.

Sapir, Edward. 1921. *Language*. New York: Harcourt, Brace and World.

Schlegel, August Wilhelm. 1818. *Observations sur la langue et la litérature provençales*. Paris: Librairie grecque-latine-allemande.

Schlegel, Friedrich. 1808. *Über die Sprache und Weisheit der Indier*. Heidelberg: Mohr und Zimmer.

Schleicher, August. 1859. Zur Morfologie der Sprache. *Mémoires de l'Académie Impériale des Sciences de St. Petersbourg*, ser 4, 1(7): 1–38.

Schultze, Wolfgang M. 2007. Review of *The world atlas of language structures*, edited by Martin Haspelmath, Matthew S. Dryer, David Gil and Bernard Comrie. *Studies in Language* 31: 445–63.

Siewerska, Anna. 1984. *The passive: A comparative linguistic analysis*. London: Croom Helm.

Silverstein, Michael. 1976. Hierarchy of features and ergativity. In R. M. W. Dixon (ed.), *Grammatical Categories in Australian Languages*, pp. 112–71. Canberra: Australian Institute of Aboriginal Studies.

Smith-Stark, T. Cedric. 1974. The plurality split. *Papers from the Annual Regional Meeting of the Chicago Linguistic Society* 10: 657–71.

Smythe, W. E. 1948/9. *Elementary grammar of the Gumbaiŋgar language (north coast, N. S. W.)*. (Oceania Monograph 8). Sydney: Australian National Research Council.

Storch, Anne. 2013. Knowing, smelling and telling tales in Luwo. In Alexandra Y. Aikenvald and Anne Storch (eds.), *Perception and cognition in language and culture*, pp. 47–68. Leiden: Brill.

Tida, Syuntarô. 2006. A grammar of the Dom language, a Papuan language of Papua New Guinea. PhD thesis, University of Kyoto.

Tonhauser, Judith. 2007. Nominal tense? The meaning of Guaraní nominal temporal markers. *Language* 83: 831–69.

2008. Defining cross-linguistic categories: The case of nominal tense. Reply to Nordlinger and Sadler (2008). *Language* 84: 332–42.

Uusküla, Mari and Urmas Sutrop. 2010. The puzzle of two terms for red in Hungarian. In Jan Wohlgemuth and Michael Cysouw (eds.), *Rara and rarissima: Documenting the fringes of linguistic diversity*, pp. 359–76. Berlin: De Gruyter Mouton.

Vennemann, Theo. 1974. Theoretical word order studies: Results and problems. *Papiere zur Linguistik* 7: 5–25.

Zeshan, Ulrike. 2004. Interrogative constructions in signed languages: Cross-linguistic perspectives. *Language* 80: 7–39.