

OXFORD
LINGUISTICS

Ian
Roberts

**Parameter
Hierarchies
& Universal
Grammar**

Parameter Hierarchies and Universal Grammar

RETHINKING COMPARATIVE SYNTAX

PUBLISHED

The Syntax of Yes and No

Anders Holmberg

Verb Placement in Romance

A Comparative Study

Norma Schifano

Person, Case, and Agreement

The Morphosyntax of Inverse Agreement and Global Case Splits

András Bárány

Parameter Hierarchies and Universal Grammar

Ian Roberts

IN PREPARATION

Case and Agreement Alignment in Romance and Beyond

Michelle Sheehan

A Featural Typology of Bantu Agreement and Nominal Licensing

Jenneke van der Wal

Rethinking Verb Second

Edited by Rebecca Woods and Sam Wolfe

Parameter Hierarchies and Universal Grammar

IAN ROBERTS

OXFORD
UNIVERSITY PRESS

OXFORD

UNIVERSITY PRESS

Great Clarendon Street, Oxford, OX2 6DP,
United Kingdom

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide. Oxford is a registered trade mark of
Oxford University Press in the UK and in certain other countries

© Ian Roberts 2019

The moral rights of the author have been asserted
First Edition published in 2019

Impression: 1

All rights reserved. No part of this publication may be reproduced, stored in
a retrieval system, or transmitted, in any form or by any means, without the
prior permission in writing of Oxford University Press, or as expressly permitted
by law, by licence or under terms agreed with the appropriate reprographics
rights organization. Enquiries concerning reproduction outside the scope of the
above should be sent to the Rights Department, Oxford University Press, at the
address above

You must not circulate this work in any other form
and you must impose this same condition on any acquirer

Published in the United States of America by Oxford University Press
198 Madison Avenue, New York, NY 10016, United States of America

British Library Cataloguing in Publication Data
Data available

Library of Congress Control Number: 2019931158

ISBN 978-0-19-880463-5

Printed and bound by
CPI Group (UK) Ltd, Croydon, CR0 4YY

Links to third party websites are provided by Oxford in good faith and
for information only. Oxford disclaims any responsibility for the materials
contained in any third party website referenced in this work.

*This book is dedicated to Undeg Enfys Moody, of Cardiff, and
to the memory of 'Thomas Roberts', allegedly of Llandudno.*

Contents

<i>Acknowledgements</i>	xi
<i>List of abbreviations</i>	xiii
Introduction	1
1. Parameters	11
1.1 The ‘classical’ view of Principles and Parameters and its problems	11
1.1.1 Chomsky (1981) and its antecedents	13
1.1.2 The explanatory value of the P&P approach	16
1.1.3 The scope of P&P theory	18
1.2 Questions for classical parameter theory: the Romance languages	23
1.2.1 Variation in subject clitics	24
1.2.2 Variation in negation	31
1.2.3 Variation in enclisis of object clitics	34
1.2.4 Variation in past-participle agreement	39
1.2.5 Variation in the nature and choice of aspectual auxiliaries	41
1.2.6 A comparison with Japanese	46
1.3 Micro- and macroparameters	53
1.3.1 In favour of microparameters	58
1.3.1.1 Formal features and the Borer–Chomsky Conjecture	58
1.3.1.2 Microparameters and restrictiveness	61
1.3.1.3 Language acquisition	65
1.3.1.4 Underspecification and the form of parameters	68
1.3.1.5 Conclusion on microparametric approaches	71
1.3.2 Macroparameters: Baker (2008a)	72
1.4 Combining micro- and macroparameters	75
1.4.1 A taxonomy of parameters	75
1.4.2 Macroparameters	77
1.4.3 Mesoparameters	78
1.4.4 Microparameters	80
1.4.5 Nanoparameters	85
1.4.6 Conclusion	88
1.5 Parameter hierarchies	89
1.5.1 Parameters and the three factors of language design	90
1.5.2 Two third-factor principles	92
1.5.3 The learning procedure	94
1.5.4 Consequences of the approach: a parameter hierarchy	97
1.5.5 Extending emergentism	99
1.6 Summary and conclusion	101

2. Word order and the Final-Over-Final Condition	103
2.1 Introduction: the head parameter, antisymmetry, and linearization	103
2.2 The Final-Over-Final Condition	111
2.3 Empirical motivation for FOFC	113
2.3.1 Clausal word order in Germanic	113
2.3.2 FOFC at the CP/TP level	117
2.3.3 Mixed projections in the nominal domain: Finnish and Latin	122
2.3.4 Morphology	126
2.3.5 Diachronic evidence	131
2.3.6 Conclusion: summary of the empirical motivation for FOFC	137
2.4 Accounting for FOFC	138
2.4.1 FOFC in terms of movement: Biberauer, Holmberg, & Roberts (2014)	141
2.4.2 Extended projections and locality	145
2.4.3 FOFC and approaches to linearization	148
2.4.4 Labelling, cyclicity, and FOFC	156
2.5 The theory of word-order variation	164
2.5.1 Formulating the word-order hierarchy	167
2.5.2 FOFC, Universal 20, and the word-order hierarchy	177
2.5.3 The word-order hierarchy and roll-up parameters	182
2.6 Conclusion: interacting factors in cross-linguistic word-order variation	187
3. Null subjects	191
3.1 Introduction	191
3.2 Preliminaries: a brief history and initial typology of null subjects	192
3.2.1 A brief history of the null-subject parameter	192
3.2.2 Consistent NSLs	199
3.2.3 Partial NSLs	207
3.2.4 Radical/discourse pro-drop languages	216
3.2.5 A typology of NSLs (first pass)	223
3.3 The internal structure of ' <i>pro</i> '	228
3.3.1 Barbosa (to appear): similarities between partial and radical NSLs	228
3.3.2 Extension to CNSLs	232
3.3.3 The typology of NSLs (second pass)	237
3.4 The nature of ' <i>pro</i> '	238
3.4.1 Aspects of the semantics of ' <i>pro</i> '	239
3.4.2 ' <i>Pro</i> ' and the Person feature	243
3.4.3 ' <i>Pro</i> -licensing' and the typology of null subjects	247
3.5 Arbitrary pronouns	254
3.5.1 Arbitrary ' <i>pro</i> ' in CNSLs and the external argument	255
3.5.2 Interpretations of arbs	261
3.5.3 Four observations about English arb elements	265

3.5.4 Arb across languages	269
3.5.4.1 The Finnish G-pronoun (Holmberg 2010b)	269
3.5.4.2 German impersonal passives	271
3.5.4.3 Turkish impersonal passives	275
3.5.5 Two ways to license arbitrary pronouns	278
3.6 φ -features: hierarchies and interfaces	281
3.6.1 φ -features, φ -hierarchies, and φ -parameters	283
3.6.2 The question of 'rich agreement'	289
3.6.3 No-choice parameters and φ -features	293
3.7 Conclusions: the nature of null subjects and the null-subject parameter hierarchy	295
4. Incorporation	301
4.1 Introduction	301
4.2 The internal structure of and relations in DP	302
4.2.1 N-D relations: Longobardi (2008)	302
4.2.1.1 Longobardi (2008) on nominal mapping	302
4.2.1.2 Parameters of nominal mapping	306
4.2.2 Person and the internal structure of DP	310
4.2.3 Conclusion	314
4.3 Polysynthesis and noun-incorporation	315
4.3.1 Baker (1996): the basic generalization and its consequences	315
4.3.2 Branigan (2012): multiple head-movement	322
4.4 Syntactic analyticity	326
4.4.1 Analyticity in DP	326
4.4.2 Analyticity in vP	330
4.4.3 Analyticity in operators	330
4.4.4 Analytic modification	332
4.4.5 Conclusion	334
4.5 Conclusions: a hierarchy for individual-denoting features	337
5. Verb-movement	342
5.1 Introduction	342
5.2 A very brief excursus into Davidsonian event semantics (Davidson 1967)	343
5.3 Tense and the event variable	347
5.3.1 Strong-Tense languages: verb-movement in Romance	348
5.3.2 Weak-Tense languages: English and elsewhere	359
5.3.3 The English auxiliary system	371
5.4 A No-Tense language: Chinese again	378
5.5 Further variation in verb-movement	383
5.5.1 V-initial languages	383
5.5.1.1 Type A V-initial languages	384
5.5.1.2 Type B V-initial languages	394
5.5.2 Germanic verb-second	398
5.6 Diachronic implications of parameter taxonomy: conditional inversion in the history of English	403
5.7 Conclusions: macrotypology of Person and Tense	409

6. Case and alignment	419
6.1 Introduction	419
6.2 Passives	420
6.2.1 Adjectival and verbal passives; passives and impersonals	420
6.2.2 Voice, passives, and external θ -roles	430
6.2.3 A parameter hierarchy for passives	439
6.2.3.1 The core parameter	440
6.2.3.2 The generalization parameter	441
6.2.3.3 The restriction parameter	444
6.2.3.4 The EPP parameter	446
6.2.3.5 Suppression of Voice's Person features	449
6.2.4 Some putative universals of passives	453
6.2.5 Further questions: passives of verbs with clausal complements	458
6.2.6 Conclusion	464
6.3 Ergativity	464
6.3.1 Introduction	464
6.3.2 The nature of ergative alignment	467
6.3.3 A parameter hierarchy for ergative alignments	472
6.3.4 Some implicational universals concerning ergativity	481
6.3.5 Conclusion	486
6.4 Other Case-related hierarchies: causatives and ditransitives	487
6.4.1 Causatives	487
6.4.2 Ditransitives	499
6.4.3 Conclusion: the nature of alignment hierarchies	506
6.5 Conclusion: Case macroparameters	516
7. Wh-movement and negation	528
7.1 Introduction	528
7.2 Wh-parameters	529
7.2.1 Wh-movement parameters	529
7.2.2 Q-particles	547
7.2.2.1 Cheng's (1991) generalizations	548
7.2.2.2 Abstract Q-particles	553
7.2.3 Conclusion	563
7.3 Negation	563
7.3.1 Negative concord and multiple negation	563
7.3.2 Clausal negators	578
7.3.3 Conclusion	599
7.4 Conclusion	600
8. Conclusion: towards a minimalist theory of syntactic variation	605
<i>References</i>	621
<i>Index of languages</i>	685
<i>Index of names</i>	694
<i>Subject index</i>	703

Acknowledgements

My thanks are due first and foremost to the European Research Council (ERC), whose award of the Advanced Grant ‘Rethinking Comparative Syntax’ (ReCoS, ERC AdG No. 269752) in the period 2011–17 funded the research reported here, along with much else (see <https://recos-dtal.mml.cam.ac.uk/>). I am grateful also to the University of Cambridge for awarding me sabbatical leave in the period October 2015 to January 2018, in which the first drafts of what follows were written.

Many people have contributed in different ways to the development of the ideas that I have tried to elucidate here, although nobody other than me should be held responsible for the end result. Above all, my thanks to the original three (Principal/Senior) Research Associates on the ReCoS project: Theresa Biberauer, Michelle Sheehan, and Jenneke van der Wal. I was the luckiest PI in history. Additional thanks to the other Recosians: András Bárány, Tim Bazalgette, Alison Biggs, Jamie Douglas, Georg Höhn, Anders Holmberg, Kari Kinn, Marieke Meelen, and Sam Wolfe. And to the extended Recosians: Víctor Acedo-Matellán, Hannah Ajer, Luigi Andriani, Alastair Appleton, Jim Baker, Ekaterina Chernova, Valentina Colasanti, Alice Corr, Maia Duguine, Anna Hollingsworth, Freddy Xuhui Hu, Aritz Irurtzun, Cherry Lam, Adam Ledgeway, Pino Longobardi, Marios Mavrogiorgos, Dimitris Michelioudakis, Moreno Mitrović, Iain Mobbs, Joe Perry, Afra Pujol i Campeny, Craig Sailor, Norma Schifano, Giuseppina Silvestri, Ioanna Sitaridou, Julio Chencheng Song, Maria Olimpia Squillaci, Ianthi Tsimpli, George Tsoulas, Bert Vaux, Sten Vikner, George Walkden, Jeffrey Watumull, and David Willis.

While on leave I spent almost three semesters at the University of Connecticut. I’d like to thank the faculty and students there for making me so welcome and effectively making UConn my second home: Karina Bertolino, Jonathan Bobaljik, Xenia Bogomolets, Željko Bošković, Andrea Calabrese, Pietro Cerrone, Christos Christopoulos, Marcin Dadan, Paula Fenger, Yoshiki Fujiwara, Jon Gajewski, Magda Kaufmann, Stefan Kaufmann, Pasha Koval, Lily Kwok, Renato Lacerda, Sabine Laszakovits, Diane Lillo-Martin, Gabriel Martinez Vera, Troy Messink, Jairo Nunes, Hiromune Oda, Vanessa Petroj, Roberto Petrosino, Hiroaki Saito, Mamoru Saito, Laura Snider, William Snyder, Jon Sprouse, Adrian Stegovec, Brendan Sugrue, Yuta Tatsumi, Harry van der Hulst, Shuyan Wang, Susi Wurmbrand, Chantale Yunt, and Shen Zheng. Special thanks to Andrea Calabrese and Jairo Nunes for co-teaching and to Tamara Cohen for making everything happen.

When I wasn’t in Storrs, CT, I was mostly in Cambridge, MA (aka the Other Cambridge, New Cambridge, Newtowne, etc.). There, thanks to the good offices of Jim Huang, I was able to work in Harvard’s Widener Library, where the majority of the pages to follow were written. Thanks to Jim, and to Bob Berwick, Gennaro Chierchia, Noam and Valeria Chomsky, Tolli Eythórsson, Jay Jasanoff, Loes Koring, Shigeiru Miyagawa, and Jenneke van der Wal for dinners and conversations in and around that Cambridge.

And then to my first linguistic home, the University of Southern California, where I taught in Fall 2017. Thanks to Thomas Borer, Bhamati Dash, Hajime Hoji, Stefan Keine, Audrey Li, Mythili Menon, Roumi Pancheva, Daniel Plesniak, Barry Schein, Andrew Simpson, and Maria-Luisa Zubizarreta. Across town at UCLA, thanks to Tim Stowell (especially for a quick outdoor tutorial on the structural representation of tense), Anoop Mahajan, Hilda Koopman, and Dominique Sportiche. Last but not least to my wonderful hosts in West Hollywood, Ben and Javier.

Many others have contributed in different ways to the ideas to follow. It's unrealistic to try to thank everyone, and so my apologies to those I may have missed. But anyway, for advice, comments, encouragement, help with fact-checking, theory-checking, brain-checking, reality-checking, and feature-checking, as well as all the other elements that go into creating a thing of this kind, my thanks to, at least: Edith Aldridge, Laura Arman, Pilar Barbosa, Balthasar Bickel, Phil Branigan, Sonia Cyrino, Probal Dasgupta, Michel deGraff, Stéphanie Durrelman-Tame, Luis Eguren, Yasuyuki Fukatomi, Roland Hinterhölzl, Marit Julien, Sahoo Kalyamalini, Mary Kato, Richie Kayne, Jaklin Kornfilt, Diane Massam, Jim McCloskey, Neil Myler, Magda Oiry, Luigi Rizzi, Ur Shlonsky, Gary Thoms, Raffaella Zanuttini, and Hedde Zeijlstra.

Last but not least, a special thanks to Fritz Newmeyer. Fritz's 2005 book *Possible and Probable Languages* 'roused me from my dogmatic slumbers' regarding the classical principles-and-parameters approach. Although, as will be evident, I heartily disagree with Fritz's conclusions, I felt I had to justify my disagreement. At length. I'm sure Fritz in his turn will disagree with much—if not all—of what I say here. And that is exactly as it should be. Ideas are not houseplants, as wiser heads have observed.

Finally (really!), thanks to Julia Steer and Vicki Sunter at OUP, who don't count words (I hope).

Ian Roberts

(Old) Cambridge
August 2018

List of abbreviations

In this book, I cite examples from a wide range of sources on various languages whose glossing conventions are not always consistent. In the interests of consistency, I have tried to make the glosses as uniform as possible, although I have not tried to follow the Leipzig conventions. My apologies to the authors of the sources I have used if they feel that I have in any way misrepresented their data. The glosses are as uniform and accurate as possible, although in certain cases either uniformity or accuracy (or both) may have been lost owing to lack of information or simply ignorance on my part. I am responsible for all errors and inaccuracies of this kind. When in doubt, the reader is advised to consult the original source.

&(P)	coordination (phrase)
1/2/3	first/second/third person
=	cliticized to
.	affixed to/fused with
A	Agent/Actor, namely subject of a transitive clause
Abl	ablative case/Case
Abs	absolutive case/Case
Acc	accusative case/Case
AcI	accusative with the infinitive
Adj	adjunct
Adv	(i) adverb (ii) adverbial case
Af	affix
Agr	agreement
AgrOP	object agreement phrase
AgrP	agreement phrase
Ant	anterior
Aor	aorist
AP	adjectival phrase
A–P	articulatory–perceptual (interface)
Appl(P)	applicative (phrase)
Arb	arbitrary pronoun/reference
Asp	aspect(ual)
AspP	aspect phrase
ATB	across the board
Aux(P)	auxiliary (phrase)
AuxV	auxiliary–verb order

BCC	Borer–Chomsky Conjecture
BP	Brazilian Portuguese
C	head of CP and/or complementizer position
Caus	causative
CI	conditional inversion
C–I	conceptual–intentional (interface)
CL	clitic
Cl	classifier
CLLD	clitic left dislocation
CNSL	consistent null-subject language
COMP	complementizer
Compl	completive aspect
Cond	conditional (mood)
Conj	conjunction
Cont	continuous
Cop	copular
CP	complementizer phrase
Dat	dative case/Case
Decl	declarative
Def	definiteness
Dem(P)	demonstrative (phrase)
Distr	distributive
DM	Distributed Morphology
DO	direct object
DOC	double object construction
DOM	Differential Object Marking
D(P)	determiner (phrase)
Du	dual
ECM	Exceptional Case Marking
ECP	Empty Category Principle
EF	edge feature
ENE	Early Modern English
EP	European Portuguese
EPP	Extended Projection Principle
Erg	ergative (case)
EST	Extended Standard Theory
Expl	expletive

F	(i) feature (ii) feminine
FE	Feature Economy
FF	formal feature
Fin(P)	finiteness (phrase)
FL	language faculty
FLB	Faculty of Language in the Broad sense
FLN	Faculty of Language in the Narrow sense
Foc(P)	focus (phrase)
FOFC	Final-Over-Final Condition
Force(P)	force (phrase)
FP	functional projection
Fr	French
FRM	featural relativized minimality
Fut	future
GA	genetic algorithm
GB	Government-Binding (Theory)
Gen	genitive case/Case
Ger	gerund
H	head
Hab	habitual
HC	Haitian Creole
HOG	Holmberg's other generalization
Hon	honorific
iF	interpretable feature
IE	Indo-European
IG	Input Generalization
Imp	imperative
Ind	indicative
Indef	indefinite
Ine	inessive
Inf	infinitive
I(NFL)	(verbal) inflection
Inst	instrumental case/Case
Int	interrogative
Intr	intransitive
Inv	inversion

IO	indirect object
IOCL	indirect-object clitic
IP	inflectional phrase
Ipfv	imperfective
It	Italian
KP	case phrase
L ₁	first language, native language
L ₂	second language
L2A	Second Language Acquisition
LAD	Language Acquisition Device
LCA	Linear Correspondence Axiom
LF	Logical Form
LI	lexical item
Loc	locative case/Case
LocCL	locative clitic
M	masculine
ME	Middle English
MGP	Modularized Global Parametrization
MHG	Middle High German
Mod	Modern
Mod(P)	modality (phrase)
MMM	Maximize Minimal Means
MP	Minimalist Program
N	neuter
NE	Modern English
Neg	negator/negative
NegP	negation phrase
NI	noun incorporation
NID	Northern Italian dialect
Nmlz	nominalizer
Nom	nominative
N(P)	noun (phrase)
NPI	negative polarity item
NSL	null-subject language
Num(P)	number (phrase)
O	(i) old (ii) object
Obl	oblique

OC	Old Chinese
OCL	object clitic
OCS	Old Church Slavonic
OE	Old English
OHG	Old High German
ON	Old Norse
Op	operator
Opt	optative
OV	object–verb (order)
OVS	object–verb–subject (order)
P	(i) object of a transitive clause (ii) parameter/property
P&P	Principles and Parameters
Pass	passive
Pat	patient
PCM	Parametric Comparison Method
Perf	perfect
Pers	person
PF	Phonological Form
Pfv	perfect(ive)
PIC	Phase Impenetrability Condition
(P)IE	(Proto-)Indo-European
Pl	plural
PLD	Primary Linguistic Data
PNSL	partial null-subject language
Pol(P)	polarity (phrase)
Poss	possessive
P(P)	preposition(al phrase)
ppt	past/passive participle
Pred	predicate
Pres	present
Pret	preterite
Prt(P)	particle (phrase)
Pst	past
Q	question/interrogative (particle)
Q(P)	quantifier (phrase)
QUD	question under discussion
RAH	rich agreement hypothesis

ReCoS	Rethinking Comparative Syntax
Refl	reflexive
Rel	relative/relativizer
RM	relativized minimality
RNSL	radical null-subject language
Ro	Romanian
S	(i) subject of an intransitive clause (ii) sentence (iii) subject
SAI	Subject–Auxiliary Inversion
SATBG	Start At The Bottom generalization
SC	small clause
SCL	subject clitic
SF	stylistic fronting
Sg	singular
Sic	Sicilian
SID	Southern Italian dialect
SMT	Strong Minimalist Thesis
SO	syntactic object
SOV	subject–object–verb (order)
Sp	Spanish
Spec	specifier position
Stat	stative
Sub	subordinate
Subjunc	subjunctive
Suf	suffix
SVO	subject–verb–object (order)
t	trace (of moved element)
TLA	Triggering Learning Algorithm
TMA	Tense, Mood, and Aspect
T(P)	tense (phrase)
Tr	transitive
uF	uninterpretable feature
U ₂₀	Universal 20
UG	Universal Grammar
UP	Uniformitarian Principle
V ₁	verb-initial
V ₂	verb-second

v_i	feature value
VO	verb–object (order)
VOS	verb–object–subject (order)
v(P)	light verb (phase)
V(P)	verb (phrase)
VSO	verb–subject–object (order)
WALS	<i>World Atlas of Language Structures</i>
X(P)	any category/maximal projection

Introduction

This book is an extended reflection on the nature of morphosyntactic variation in natural language. Probably one of the easiest things to observe about language is its variability: languages and dialects vary over time and across space. This variation is readily apparent even to the most casual observer. Moreover, all aspects of the structure of language seem to be open to variation: phonology, syntax, and morphology, as well as the lexicon. Linguistic variants are culturally sanctioned as languages in the everyday sense of the term, and so we speak of the English language, the French language, and so on. In the terminology of Chomsky (1986b), these cultural entities are forms of E-language, but individuals we identify as English speakers, French speakers, etc. have internalized variant I-languages. So variation is found at both the E-language and the I-language level.

At the same time, the search for language universals has been an abiding concern for linguists and linguistic theory. In recent decades, this has taken two principal forms. On the one hand, the field of language typology has sought to observe language universals of one kind or another by directly cataloguing common structural features across many languages. This approach was initiated by Greenberg (1963/2007), with a sample of thirty languages. At the time of writing (early 2018), *The World Atlas of Language Structures* (WALS henceforth, available online at <http://wals.info>) reports data from a total of 2,679 languages. Strikingly, while Greenberg cited 45 putative universals in his original paper, many working in this field now feel that the notion of universal may be chimerical (see, e.g., Evans & Levinson 2009; Bickel 2014; 2015). Nonetheless, many universals have been proposed: the Universals Archive at the University of Konstanz lists over 2,000 (see <http://typo.uni-konstanz.de/archive/>).

The other form of investigation of language universals in recent decades has been directly inspired by the work of Noam Chomsky (see in particular Chomsky 1965; 1975; 1980; 1986b). Chomsky argues that there must be a biological predisposition to language. Again, the argument is based on two observations about language which are readily made: first, that language is an extremely complex phenomenon, and, second, that young children acquire their native language with apparent speed and ease. Together, these two observations lead to the conclusion that there must be some inbuilt cognitive bias which facilitates language acquisition by constraining the hypothesis space within which language learning can operate. The most direct—although certainly not the only—way to guarantee this is by constraining the form of a possible grammar of a human language, i.e. defining the class of possible I-languages. This amounts to constructing a theory of possible human grammars (taken to be a subset of the set of

grammars, and therefore I-languages). Such a theory is universal by definition; hence the constraints on possible grammars/I-languages should manifest themselves as structural universals of language. This set of constraints is generally referred to as Universal Grammar (UG henceforth); UG is thus ‘the general theory of I-languages’ (Berwick & Chomsky 2016: 90). The notion of ‘inbuilt cognitive bias’ just alluded to is often thought to be genetically determined; if so, the possible form of grammars represents a modern version of the notion of ‘innate ideas’. Put simply, UG is innate.

In what follows I will adopt a Chomskyan perspective, in that I accept the premise that some constraint on the possible form of natural-language grammars is required in order to account for the two observations just described. How directly such constraints need to be genetically coded is, however, a question which I will not decide in advance; this is an overarching issue which I will not directly address in what follows, although I will return to it briefly in Chapter 8.

Adopting a Chomskyan perspective entails the postulation of universals, as we have seen: UG defines the general form of a possible grammar of a human language. In this context, the simple observation of massive structural variation, at all levels and in all observable times and places, raises a problem. How can we reconcile such easily observed, culturally sanctioned linguistic diversity with the fact that human linguistic competence appears to be a readily acquired cognitive capacity? Moreover, human linguistic competence appears to be *uniquely* human, and our species is known to be genetically rather homogeneous (see, e.g., Reed & Tishkoff 2006). Independently of whether one assumes some form of biological predisposition to language of the Chomskyan kind (but all the more so if one does), reconciling the attested linguistic diversity with the cognitive and genetic unity of the human species is a non-trivial matter. At its most general, this is the question this book tries to address.

Since the early 1980s, mainstream generative grammar has developed an approach which addresses this question by postulating that UG allows for variation: this is known as the principles and parameters approach, first articulated in detail by Chomsky (1981) (P&P henceforth). The principles of UG articulate the invariant constraints on grammatical form motivated by the twin observations of linguistic complexity and apparent ease of language acquisition. Allied to these principles, at least as was first thought, are parameters specifying a restricted range of variation. Hence both what varies and what is invariant were seen as part of the innate linguistic endowment.

It was almost immediately realized that the P&P approach could create a link between the Greenbergian and Chomskyan approaches to linguistic variation and diversity. The parameters of UG may underlie, directly or indirectly, the typological variation and diversity observed in the Greenbergian tradition. This central idea has led to a great deal of productive cross-linguistic research, and is what motivates the present work. In the remainder of this Introduction, I want to illustrate how the Greenbergian and Chomskyan traditions have interacted in one particular domain, that of cross-linguistic word-order variation, showing both the advantages and disadvantages of standard P&P and sketching a novel general approach of a kind which constitutes the central idea in this book.

Among the 45 universals observed in Greenberg's classic 1963/2007 paper, there were several which dealt with word order. Of these, as shown by Dryer (1992) in his reassessment of Greenberg's early observations in the light of improved language sampling (intended among other things to remove the Indo-European bias in Greenberg's sample) and a database of 625 languages, a number fail to stand up but, perhaps more interestingly, a good number still do hold. Consider as an illustration the correlation between VO order and prepositions and OV order and postpositions (Greenberg's Universals 3 and 4; Universal 3 was originally restricted to VSO languages, but this was later extended to SVO languages by W. Lehmann 1973; Vennemann 1974; and Hawkins 1983). The figures from the latest version of *WALS* are as follows (these figures exclude inpositions and cases of 'no dominant order' in the interests of exposition):

(1)	OV & Postpositions	472	
	OV & Prepositions	14	
	VO & Postpositions	41	
	VO & Prepositions	454	(Dryer 2013a; 2013b)

Here there is a total of 981 languages, and so the figures approximate closely to percentages. Hence more than 90% of the languages sampled show the correlation. In terms of standard P&P theory, this correlation is captured by the Head Parameter, which we can formulate as in (2):

- (2) In X' , X {precedes/follows} its complement YP .

(See Stowell 1981; Huang 1982a; Koopman 1984; Travis 1984 for early formulations of this idea, differing in detail from this one; this formulation is from Rizzi 1988.) This parameter exploits the category-neutral nature of the X' -schema in order to state the cross-categorical generalization, a possibility first adumbrated by Lightfoot (1979) and developed in Hawkins (1983); see Roberts (2017a: 31–3). Here we see the P&P idea in action: the X' -theoretic notions of 'head', 'complement', and ' X ' are defined by UG, and are facets of invariant principles of phrase structure. The options 'precede' or 'follow' represent the parameter, also stated at the level of UG, and are also therefore taken to be part of the innate endowment.

But we can also see a problem: a minority of languages appear to disobey the correlation. In this connection, there are several possibilities. We should, of course, check that the reported information is accurate. Assuming it is, then we need to look closely at the languages in question and see whether a plausible analysis which would bring them into line with (2) is possible. Since *WALS* is based on reported surface facts, it is in principle always possible that a generative analysis—referring to a deeper level of syntactic analysis—may be available, which will solve the problem. It is reasonable to think that this is possible in at least some cases, but it is also likely that some of these 55 languages will remain problematic. Of course, the non-surface-based analysis can lead to the opposite situation too. For example, German, Dutch, and Frisian are reported in *WALS* as combining 'no dominant order' for OV/VO with prepositions (no data on Afrikaans is reported for these

features in *WALS*, but it well-known that Afrikaans is broadly very similar to Dutch in these respects: see Biberauer 2003). German is stated to be a language ‘in which word order is primarily determined syntactically, but in which there are competing OV and VO constructions’ (Dryer 2013a). In the generative tradition since Koster (1975) these languages have been analysed as underlyingly OV, with surface VO orders derived by verb-movement to the left of the object in the relevant contexts (main clauses where there is no auxiliary). Since these languages are prepositional (in fact there are also some postpositions, but I will leave this complication aside here), this analysis adds up to four further languages in the OV & Preposition group.

It is very likely, then, that we are faced with real counterexamples to the prevailing tendency, even if some of the 55 problematic languages could be shown either to have been misreported or to be amenable to a plausible ‘deeper’ analysis which assigns them to one of the other categories.¹ Since parameters are taken as ‘hard’, UG-given constraints, we cannot simply replace (2) with a tendency, or attach some kind of weighting or preference to it (at least not in a direct way). The only option available, in that case, is to make (2) less general, and relativize it to categories. Let us then replace (2) with (3):

- (3) a. In V' , V {precedes/follows} its complement.
- b. In P' , P {precedes/follows} its complement.

We can now readily capture the data summarized in (1). The 14 OV, Prepositional languages set (3a) to ‘follow’ and (3b) to ‘precede’ (as do Afrikaans, Dutch, German, and Frisian) and the 41 VO, Postpositional languages do the opposite.

It is clear, though, that such empirical adequacy has a heavy theoretical price. We now have no way to capture the overwhelming tendency towards harmonic orders in parametric terms (of course, we could introduce some other means, or appeal to psycholinguistic factors following proposals such as those in Hawkins 1983; 1994; 2004). The figures in (1) could just as easily have been the other way around, if the parameters regulating word order are those in (3). Moreover, this exercise can be repeated for all heads and complements, giving us as many head parameters as there are heads. This is clearly not a good situation.

It is here that one of the central ideas in what follows comes into play. In essence, in order to capture both the fact that there are exceptions to (2) and the fact that cross-categorical harmony is strongly preferred, we need to able *both* to state the parameter in category-neutral terms as in (2) *and* to weaken it to specific

¹ Sheehan (2017c: 89) draws attention to the fact that three times as many languages combine VO with Postpositions as combine OV with Prepositions. As she points out, only the latter order violates the Final-Over-Final Condition (FOFC; see (9) below and the extensive discussion in Ch. 2). This order is also confined to fewer macro-areas, language families, and genera than the inverse non-harmonic order. In fact, the order $P > DP > V$ appears to be confined to Indo-European (non-English West Germanic, Persian, Tajik, Kurdish, and Sorbian) and one Semitic language, Tigré. This geographical and genetic rarity does not alter the fact that the order is both disharmonic and non-FOFC-compliant, of course, but it is consistent with the notion that it constitutes an exception to prevailing tendencies.

categories as necessary, along the lines of (3). This can be achieved if we assume (i) that all parameters are properties of individual heads (as in (3)), but (ii) that there are Independent Factors (IF) which can cause groups of heads—over which we can generalize with an appropriately formulated feature system—to act in concert. We must further assume that (iii) there is a preference for the IFs to act in this way, but that this preference is ‘soft’, i.e. defeasible. If we make these assumptions, and if we can understand fully the weighting of the preference, then we can capture the data in (1) and maintain that ‘with overwhelmingly greater than chance frequency’ (in Greenberg’s apposite formulation) the category-neutral parameter in (2) holds.

It is very important to see that it follows from (i–iii) in the preceding paragraph that (2) cannot be a parameter of UG in the classical sense. Instead, it is an epiphenomenon: the result of the interaction of a much more fragmented set of parameters each relating to individual heads, the IFs, and the preference for IFs to cause heads to act in concert alluded to above. Spelling out how these properties interact in detail across a range of cross-linguistically variant phenomena is one of the central goals of this book.

Let us now be more precise, and begin to flesh out the basic ideas just set out. One of the things we must articulate is what the feature system which underlies parametric variation is. Following Chomsky (1995: 6), we propose the notion of the ‘fragmented set of parameters’ just alluded to as the formal features of functional heads. For concreteness, let us assume that there is a privative feature determining head–complement linearization—call it [F] (which perhaps stands for Follow). Hence, if a head *H* has [F], it follows its complement; otherwise *H* precedes its complement.

What are the IFs that cause classes of *H*[F]s to act together? Following earlier work by Roberts & Roussou (2003: 201) and Roberts (2007a: 275), I propose these are as follows:

- (4) (i) Feature Economy (FE):
Postulate as few formal features as possible.
- (ii) Input Generalization (IG):
Maximize available features.

Together, FE and IG form a search/optimization algorithm, with FE minimizing features where possible and IG maximizing detected features; Biberauer (2016; 2017a; 2018) unifies FE and IG as a single constraint, Maximize Minimal Means (MMM). Parametric variation arises from the fact that UG leaves the inventory of formal features underspecified for individual I-languages.

In the case of our example, harmonic head-initial order represents a default option, since *F* is not present, in line with Feature Economy. Input Generalization, in the case of head-initial languages, is also vacuously satisfied because *F* is entirely absent from the system. So in this case both FE and IG are compatible with the head-initial grammar because no feature is posited and the absence of the feature is maximally general. However, if the Primary Linguistic Data (PLD) the child is exposed to is not compatible with the fully head-initial grammar

(i.e. if the acquirer is led to parse input strings as head-final), then the question of the presence of F arises, IG becomes relevant, and F is assumed to be a feature of all (relevant) heads, with the consequence that a fully head-final system emerges (recall that IG and FE are both defeasible by the PLD; here FE is clearly overridden).

Categorial distinctions may also come into play. If the PLD is such that generalized head-final order cannot be maintained (i.e. if it is such that some strings must be parsed as head-final and others as head-initial), then a categorial distinction is posited as relevant for word order (with FE again overridden) and generalized (IG). Nonetheless—and this is the crucial point in relation to (2) and (3) above—no categorial distinction is needed in the encoding of linearization properties unless the PLD is ‘mixed’ in the way just described. Given ‘mixed’ PLD, the acquirer ‘redeploys’ this distinction in order to make its grammar compatible with the PLD.

We can see from the above discussion that the parametric options fall into a hierarchy. We can state the relevant hierarchy as follows:^{2,3}

- (5) a. Is F present in the system? (Y/N)
N: rigidly, harmonically head-initial language.
- b. Y: is F generalized to all (relevant) heads?
 Y: rigidly, harmonically head-final language.
- c. N: if not, is F restricted to some subset of heads, e.g. [+V], [+N]?

Option (5a) is fully consistent with FE and IG. In the case of option (b), FE is overridden by the PLD, but IG is fully obeyed. Option (5c) is the case where disharmony may arise; here a system of the kind allowed by the parameters in (3) may arise. But such categorial restrictions do not play a role as long as the acquirer takes either option (a) or option (b), and indeed may not even be present in the system at an early stage of the acquisition process (on this last point, see again Biberauer 2016; 2017a).

In addition to laying out the learning path, parameter hierarchies clearly have the capacity to encode typological and diachronic generalizations. A major result of early work on word-order typology (Greenberg 1963/2007; Hawkins 1983; Dryer 1992) is that there is a clear preference for what Hawkins refers to as Cross-Categorial Harmony. In terms of the present approach, Cross-Categorial Harmony is the consequence of the role of IG in channelling language acquisition into hierarchies as described above.

A further aspect of this approach is that the IFs in (4) are *not* taken to be part of UG, but instead represent general cognitive optimization strategies which may

² I will adopt the practice throughout of stating the parameters as yes/no questions. This is an essentially expository move, but it has the advantage of making their binary nature clear, and it reminds us that the parametric options represent ways in which the child interrogates the PLD (although this idea should, of course, not be taken too literally).

³ This hierarchy is based on earlier work (see, e.g., Roberts 2012a: 321; Biberauer, Holmberg, Roberts, & Sheehan 2014: 110), and will be significantly revised in §2.5.

well apply in other areas of cognition (on this, see especially Jaspers 2005; 2012; Biberauer 2016; 2017a; 2018). More generally the approach adopted here is in line with the three factors of language design discussed by Chomsky (2005). These are as follows (we will look at these in more detail in §1.5.2):

- (6) F1: the biological endowment, UG;
F2: experience, i.e. PLD;
F3: general optimization strategies.

The specific instantiations of the three factors relevant for parameter hierarchies are as follows:

- (7) F1: underspecification of formal features in UG;
F2: trigger experience/what the child takes up;
F3: FE and IG.

Concerning the second factor, Biberauer (2016; 2017a; 2018) points out that, for the postulation of formal features, what the child needs to observe is departures from the simplest Saussurean form–meaning mapping, e.g. multiple realization of elements, displacement, and apparent non-realization of meaningful elements (silence).

Given the three factors as in (7), we arrive at a central idea of the present work:

- (8) Parameters are emergent properties of the interaction of the three factors.

It follows from (8) that parameters are *not* directly predetermined by UG. This is a major departure from previous thinking in this area (see §1.1.1 for some historical background on parameters).

Now that we have introduced the word-order hierarchy, albeit in a rather preliminary form (see note 3), we can further illustrate how principles and parameters interact in terms of the Final-Over-Final Condition (FOFC). This is stated by Holmberg (2017a: 1) as follows:

- (9) A head-final phrase αP cannot immediately dominate a head-initial phrase βP , if α and β are members of the same extended projection.

The effect of (9) is to rule out the configuration in (10), where αP is the complement of β and γP is the complement of α , and all of α , β , γ are in the same extended projection:⁴

- (10) $*[_{\alpha P} \quad [_{\beta P} \beta \gamma P] \dots \alpha]$

⁴ Holmberg has ‘immediately dominate’ in (9), but ‘...’ in (10) allows for intervening material, i.e. further heads and complements. Since these must all be either head-final or head-initial, and there must be a ‘switch’ somewhere, (10) actually instantiates the configuration for (9) without reference to *immediate* domination, which is a more general formulation.

In other words, a head-initial category can take either a head-initial or a head-final category as its complement, while a head-final category can only take another head-final category as its complement (in the relevant domain, i.e. extended projection). The case where a head-final category takes a head-initial category as its complement corresponds to the configuration in (10), and this is ruled out.

We will consider the empirical motivation for FOFC, along with the analysis proposed in Biberauer, Holmberg, & Roberts (2014), in some detail in Chapter 2. For now let us take it for granted that (10) holds as a universal constraint (see Sheehan, Biberauer, Roberts, & Holmberg 2017 for extensive discussion and justification of this idea).

FOFC is clearly a constraint on disharmonic orders, and can be seen as a constraint on the complement-taking abilities of a final head in a disharmonic system; in fact, it expresses a dependency between heads such that lower heads are head-initial only if all higher ones are; as Biberauer (2017b: 291f.) points out, this is an instance of a wider phenomenon of ‘harmonic contiguity’, a point we will return to in several places (see §§2.4.4, 3.6.1, 4.2.3, and 6.5). As such, it automatically rules out large classes of otherwise conceivable disharmonic orders. For example, another word-order universal originally proposed by Greenberg states that auxiliaries tend to pattern with verbs, in that OV languages have the order V Aux, while VO languages have Aux V (Universal 16; Greenberg’s original formulation actually referred to VSO languages, but, as in the case of Universals 3 and 4 discussed above, we can generalize it to include SVO languages as well). Since auxiliaries are in the same extended projection as their main verb (see Grimshaw 1991, 2001; we will present and elaborate Biberauer, Holmberg & Roberts’ notion of extended projection in §2.4.2), FOFC predicts that, alongside the harmonic Aux VO and O Aux orders predicted by Greenberg’s Universal 16, we find Aux OV, where a head initial Aux P (presumably some functional category in the inflection domain of the clause) has a head-final VP as its complement, but *not* VO Aux. In VO Aux the head-final Aux P has a head-initial VP as its complement (or dominated by its complement—see note 4). This instantiates the structure in (10) (for $\beta = V$, $\alpha P = O$, and $\gamma = \text{Aux}$). Biberauer, Holmberg, & Roberts argue that this prediction is empirically correct, and we will see some of their evidence in §2.3 (see also Biberauer, Holmberg, Roberts, & Sheehan 2017).

The general consequence of FOFC in relation to the parameters for word order in (5) is that if a head H in the extended projection of a lexical category L (EP(L)) has F, then all the heads c-commanded by H in EP(L) must also have F; the properties of lower heads can be predicted on the basis of properties of higher heads (this is the ‘harmonic contiguity’ effect mentioned above; see again Biberauer 2017b: 291f.). This has the effect of eliminating a large number of logically possible disharmonic systems. Interestingly, it also has the effect of increasing the probability of harmonically head-final systems. The higher F-bearing H is in EP(L), the more options are lost. To see this, consider the heads in EP(V): on fairly standard non-cartographic assumptions about clause structure, following Chomsky (2000; 2001), V has the two options (F or not) and hence FOFC rules out nothing. At the level of v there are four options (both v and V can have F or

not), one of which, VOv, FOFC rules out. At T there are eight options, as all of T, v, and V can have F or not, four of which FOFC rules out. Finally, at the level of C the number of options available in principle doubles again, to sixteen. Of these, eleven are ruled out by FOFC. More generally, if we assign a number n to each head H in EP(L) corresponding to the number of heads in the EP H c-commands plus one, then the number of options consistent with FOFC is $n + 1$, and the number ruled out $2^n - (n + 1)$. For example, C is head 4 (since it c-commands T, v, and V) and, as just mentioned, it allows five FOFC-compliant linearizations out of a logically possible sixteen ($= 2^4$), ruling out eleven ($16 - (4 + 1)$). Hence the probability of a harmonic head-final CP is not one in sixteen, as it would be with a free choice of initial or final order at each head, but one in five. So a typological consequence of FOFC may be the relatively large number of harmonically OV languages (especially given that this is the second-choice option in the hierarchy in (5)). Regarding acquisition, it is also clear that FOFC severely restricts the set of possible disharmonic systems among which the acquirer has to search;⁵ if there are other harmonic contiguity effects of this kind such that the properties of structurally lower heads can be predicted on the basis of higher ones, then the space the acquirer needs to search is correspondingly reduced.

The interaction of FOFC with the word order in (5) is clearly of explanatory value, in that it simultaneously explains the perhaps unexpectedly large proportion of head-final systems and facilitates language acquisition by limiting the set of available options. This is in fact a case of principle (FOFC) and parameter (the word-order hierarchy) interaction of the kind that P&P theory has always attempted, not always successfully, to seek out. Of course, I do not wish to suggest that FOFC is a primitive UG principle; in Chapter 2, following Biberauer, Holmberg, & Roberts (2014) and Sheehan, Biberauer, Roberts, & Holmberg (2017), I will investigate the idea that it follows from the combination of the theories of linearization and locality/cyclicity. The word-order hierarchy emerges from the interaction of the three factors in language design, as we saw above. So, whilst this is a classic case of ‘principle’ and ‘parameter’ interaction, with the desirable consequences that follow, both the principle and the parameter derive from more elementary notions. In this way, we move towards a minimalist approach to principles and parameters, and to morphosyntactic variation in general. The prime goal of this book is to do exactly this over a range of empirical domains.

Chapter 1 looks in more depth at the motivations for rethinking the theory of comparative syntax along the lines we are proposing. Chapter 2 looks at word order and FOFC, developing and discussing a range of technical questions that were glossed over in the brief presentation above and arriving at a rather different picture of cross-linguistic word-order variation from that just sketched. Chapter 3 turns to null subjects, taking up where the discussion in Roberts & Holmberg (2010) left off, and adapting and adopting the important recent proposals in

⁵ In §§2.4.4 and 2.5, a different approach to head-initiality and head-finality will be proposed which alters many of the assumptions behind this argument. The basic force of the argument remains unchanged, however, since we continue to assume FOFC.

Barbosa (to appear). Chapter 4 looks at incorporation, extending the analysis of pronouns developed in Chapter 3 to DPs in general, and integrating the results of work on polysynthesis by Baker (1996) and Branigan (2012) as well as work on ‘deep analyticity’ by Huang (2015). In Chapter 5, I turn to verb-movement, and analyse this phenomenon along the same general lines as N-movement in the DP, developing ideas of Longobardi (1994; 2008) and Pollock (1989), and summarizing and integrating the important work on Romance by Schifano (2018), as well as work on V1 and V2 languages. Chapter 6 looks at Case and alignment parameters, building on important work on passives, ergativity, causatives, ditransitives, and applicatives by Sheehan (2017a; 2017b; in progress; Cyrino & Sheehan 2016; Sheehan & Cyrino 2018; and Sheehan & Roberts, in progress). Chapter 7 looks at parameters involving *wh*-movement and negation, in the latter case adopting and adapting the proposals in Biberauer & Zeijlstra (2012a; 2012b). Chapter 8 supplies a brief conclusion.

Parameters

1.1 The ‘classical’ view of Principles and Parameters and its problems

The P&P approach to comparative syntax has suffered some criticism in recent years; see in particular Newmeyer (2004; 2005a; 2005b), Haspelmath (2008), and Boeckx (2011; 2014; 2015; 2016). The criticisms range over both empirical and theoretical shortcomings.

The empirical criticism is quite simple: hardly any parameters have been established on which there is general agreement. In over thirty years of research, few examples have been found that generalize in a straightforward way to more than a few languages. Newmeyer (2004; 2005) makes this case based on an examination of the cross-linguistic scope of the Null Subject Parameter (as formulated in Rizzi 1982; 1986b), and concludes that the programme has proven a failure and should be abandoned.

A different line of criticism, of a more conceptual nature, has emerged from the emphasis in Chomsky’s (1995; 2005; 2007) work on reducing the content of UG to the minimum. Since the P&P model, in its ‘classical’ form as elucidated in Chomsky (1981), appears to presuppose a complex and richly structured UG, this is a problem for that model (see Newmeyer 2005a: 83). In this connection, Berwick & Chomsky (2011; 2016: 82) argue that much of the observable variation in grammatical systems reflects the nature of what they refer to as ‘the externalization process,’ i.e. the phonological/morphological interface (PF), rather than the narrow syntax (NS) itself. They suggest that true, NS-internal, syntactic variation may be negligible or nonexistent. In a different way, then, Berwick & Chomsky also advocate abandoning the earlier notion of syntactic parameter.

Boeckx (2011; 2014; 2015; 2016) echoes both Newmeyer’s and Berwick & Chomsky’s points in his critique of the standard notion of parameter. He goes on to make a distinction between the original notion of Parameter (with a capital ‘P’) and the weaker notion of ‘parameter’. The former is the concept introduced in Chomsky (1981), which, as Boeckx rightly states, is not really compatible with a minimalist system in which there are very few principles to be parametrized, and those principles which we do have (notably Merge) are unlikely to be subject to parametrization (see §1.3.1). Furthermore, he proposes the Strong Uniformity Thesis (SUT): ‘Principles of narrow syntax are not subject to parametrization; nor are they affected by lexical parameters’ (Boeckx 2015: 119); this is close to Berwick & Chomsky’s (2016) proposal that parameters are restricted to externalization processes. From this Boeckx concludes: ‘If Principles disappear, Parameters

can't be maintained' (2015: 131). Regarding the lower-case 'parameter', Boeckx states that this is a 'much watered-down notion' and that 'it is clearly devoid of any theoretical teeth, hence for me does not even begin to exist in a theoretical context' (p. 130). Boeckx critically discusses the proposals by Bošković (2008) for an NP-DP parameter and Huang's (2005; 2015) High Analyticity parameter (see §4.4 on the latter), as well as Cinque's (2013a) proposals regarding word-order parameters, seemingly on the grounds that most proposed parameters admit exceptions (i.e. his critique is effectively identical to Newmeyer's), or, in the case of Huang's High Analyticity parameter, that 'it is a continuous notion disguised as a discrete state' (Boeckx 2015: 135) (in fact Huang explicitly relates analyticity to lack of head-movement across several domains, clearly a discrete property of a grammatical system; we will look at his proposals in detail in §4.4).

Boeckx explicitly rejects what he refers to as "Greenbergizing" Plato' (2015: 135), on the grounds that language typology only relates to E-languages. This last point seems misplaced: from the perspective of generative grammar, much typological analysis seems excessively surface-oriented (just as for many typologists, generative analyses often seem too abstract), but many observations in the typological literature, including many of Greenberg's (1963/2007) original observations, can clearly tell us something about the underlying I-languages (see, e.g., recent work on Greenberg's Universal 20 by Cinque 2005 and Abels & Neeleman 2012, and §2.5.2). Surface observations may be simplistic, and may disguise deeper regularities (or possibly irregularities), but we cannot know in advance that they are not relevant for understanding natural-language grammars, and hence may—however indirectly—reflect aspects of I-language. To dismiss typological generalizations, of whatever 'depth', on a priori grounds as irrelevant to understanding the nature of possible grammatical systems seems to ignore a source of potentially useful data regarding the latter, just as to dismiss evidence from first-language acquisition (e.g. observations regarding Early Null Subjects, 'Principle-B lag', etc.) would be. We cannot know in advance of theoretically driven analysis what, if anything, a given typological observation is telling us; but to dismiss all such observations out of hand does not seem a good strategy for pursuing an understanding of the nature of cross-linguistic variation and the restrictions on possible grammatical systems. As we will see below, the fact that P&P theory brings Greenberg and Plato together, in a certain sense, gives this approach much of its power.

Boeckx further criticizes the 'Borer–Chomsky conjecture' (BCC; so named by Baker 2008a): the notion that parametric variation reduces to (a subset of) formal features of (a subset of) functional heads; see §1.3.1.1(60) for a formulation (note that this is actually consistent with SUT, as long as we hold that variation in formal features of functional heads does not determine narrow-syntactic principles, but rather interacts with them to give rise to variation). There are indeed conceptual problems with this proposal, in the absence of an adequate intensional characterization of the relevant set of features. Furthermore, since it relies on formal features as specified in the lexicon, it is radically incompatible with the general model Boeckx is arguing for. However, in §1.3.1.1 I will argue that the

BCC, although flawed, can provide the basis for a discussion of parameters which is good enough to work with.

In general, the proposals to abandon the notion of parameter altogether do not really seem to offer viable alternatives (this is not true of Berwick & Chomsky's proposal, which simply seeks to locate all the variation at one interface). For example, the alternative Newmeyer (2005a) sketches appears to be a return to the theoretical *status quo ante*, and hence brings back inductive learning of rules (this point was made by Dryer 2007a in his review of Newmeyer 2005a). Boeckx (2015: 144, table 4.1) lists a number of 'learning biases': reasoning under uncertainty (Kemp, Perfors, & Tenenbaum 2007), his own (Boeckx 2011) Superset Bias, Briscoe & Feldman's (2006) Bias/Variance Trade-off, Yang's (2002) emphasis on statistical computation and the Tolerance Principle (Yang 2016), the Elsewhere Condition, sensitivity to PF cues, and constraints on perception and memory of the kind discussed by Endress, Nespore, & Mehler (2009) and Gervain & Mehler (2010). This leads to a proposed acquisition algorithm (Boeckx 2015: 145, fig. 4.3). However, as with Newmeyer, it is not clear what Boeckx takes to be the object of acquisition; it is not parameters (or Parameters), but it is not clear what it is. If the object of acquisition reduces to rule systems (the only alternative to parameter setting yet put forward in the history of generative grammar), then, as for Newmeyer's proposals, the result is a return to the *status quo ante*. What is clear, as Boeckx (2015: 143) states, is that 'the algorithm ... needs to be developed further and, importantly, put to the test'. Since this has not yet happened, and given their highly programmatic nature, I leave these proposals aside here.

One of the principal motivations for the current work is that the P&P approach still offers a genuine prospect of reaching explanatory adequacy in the sense originally defined in Chomsky (1964), as was briefly illustrated in the Introduction; see §1.1.2 below. If, as I suggested in the Introduction, this can be done while maintaining minimalist theoretical parsimony, then we may even be able to take P&P theory beyond explanatory adequacy. For this reason, I do not propose to give up P&P theory, but it must be radically rethought. Most of this book is devoted to doing just that, but first let us consider the merits of the original approach.

1.1.1 Chomsky (1981) and its antecedents

The goal of this and the next two subsections is to recall the original scope and apparent promise of the P&P approach; hence the discussion is to a degree historical and exegetical in nature. Chomsky (1981: 3ff.) contains the first full exposition of a theory of principles and parameters. But the first use of the term 'parameter' in linguistic theory in the now-familiar sense predates Chomsky (1981) by a few years.¹ It is to be found in Chomsky (1977a: 175), which was

¹ See also Rizzi (2014: 14–17) for discussion of the earliest proposals regarding parameters, including the quoted passage from Chomsky, and in particular for discussion and illustration of the Italian/English contrasts in *wh*-movement which motivated his (1978/1982) proposal mentioned below.

originally published in 1976, based on lectures given in June 1975. It is worth quoting the relevant passage at length:

Even if conditions are language- or rule-particular, there are limits to the possible diversity of grammar. Thus, such conditions can be regarded as parameters that have to be fixed (for the language, or for particular rules, in the worst case), in language learning. We would then raise the question how the class of parameters so constituted, with rules that lack expressive power but parameters to be fixed independently, compares with the class of grammars permitted under a theory that permits articulation of conditions on application within the formulation of the rules themselves. It has often been supposed that conditions on application of rules must be quite general, even universal, to be significant, but that need not be the case if establishing a 'parametric' condition permits us to reduce substantially the class of possible rules.

We might hope to find that even if some condition C on rule application is language-particular, nevertheless some general principle determines that it applies to languages of a specific type, in which case we will have again reduced the problem, indirectly, to a property of UG.

Conditions on rule application include constraints such as subadjacency. Shortly afterwards, Rizzi (1978) argued that apparent violations of the *wh*-island condition in Italian, and hence of subadjacency, can be shown to fall under this condition (independently motivated by the evidence of Complex NP Constraint violations in Italian) if we allow that the bounding nodes for Italian are NP and S' (= CP) rather than NP and S (= TP) as is the case in English (see in particular note 25, pp. 73–4 of the 1982 reprint of this paper). It is clear that the variation between Italian and English discovered by Rizzi falls exactly under the general scenario sketched by Chomsky in the quotation above.

Chomsky (1981: 3) introduces the P&P idea as follows:

The theory of UG must meet two conditions. On the one hand, it must be compatible with the diversity of existing (indeed, possible) grammars. At the same time, UG must be sufficiently constrained and restrictive in the options it permits so as to account for the fact that each of these grammars develops in the mind on the basis of quite limited evidence...

What we expect to find, then, is a highly structured theory of UG based on a number of fundamental principles that sharply restrict the class of attainable grammars and narrowly constrain their form, but with parameters that have to be fixed by experience. If these parameters are embedded in a theory of UG that is sufficiently rich in structure, then the languages that are determined by fixing their values one way or another will appear to be quite diverse, since the consequences of one set of choices may be very different from the consequences of another set; yet at the same time, limited evidence, just sufficient to fix the parameters of UG, will determine a grammar that may be very intricate and will in general lack grounding in experience in the sense of an inductive basis.

After introducing the subsystems of UG that played a central role in GB theory (Case theory, binding theory, etc.), Chomsky (1981: 5) states: 'Ideally, we hope to find that complexes of properties differentiating otherwise similar languages are

reducible to a single parameter.' He also refers to Lightfoot (1979) for 'analogous considerations concerning language change' (1981: 6).

Concerning language acquisition, Chomsky (1981: 6) says '[e]xperience—in part, a construct based on the internal state given or already attained—serves to fix the parameters of UG, providing a core grammar, guided perhaps by a structure of preferences and implicational relations among the parameters of core theory.' He further points out (p. 11) that 'UG will provide a finite set of parameters, each with a finite number of values' and hence 'will make available only a finite class of possible core grammars'. This is followed by a discussion of the implications, or rather the lack of them, of this conclusion for formal learnability theory, since the set of grammars allowed by UG had previously been thought to be infinite.

Finally, he states: 'The grammar of a particular language can be regarded as simply a specification of values of parameters of UG, nothing more' (1981: 27). We see then that all the elements of the theory of principles and parameters of UG are introduced: those concerning language acquisition, language change, markedness, implicational relations among parameters, potential consequences for learnability theory, and, of course, the potential for significant cross-linguistic comparison. On this last point, Chomsky (p. 95) makes the connection between 'X'-parameters' and word-order typology, further linking this to language acquisition:

rules...expanding X' will appear in the particular grammar of English only as the specification that complement follows head; apart from this, the rule belongs to UG rather than the grammar of English. Universals of the sort explored by Joseph Greenberg and others have obvious relevance to determining just which properties of the lexicon have to be specified in this manner in particular grammars—or to put it in other terms, just how much must be learned as grammar develops in the course of acquisition.

Here, we see how biolinguistics meets language typology in the context of P&P theory. In a nutshell, Greenberg meets Plato.

Chomsky (1980: 67) also introduces the basic P&P idea, although rather more schematically than in Chomsky (1981). Here he makes a specific analogy with growth and speciation, saying that 'the problem of accounting for the growth of different languages... is not unlike the general problem of growth, or for that matter, speciation'. He then cites a passage from Jacob (1978):

What accounts for the difference between a butterfly and a lion, a chicken and a fly, or a worm and a whale is not their chemical components...speciation and diversification called only for different utilization of the same structural information... It is thanks to the complex regulatory circuits, which either unleash or restrain the various biochemical activities of the organism, that the genetic program is implemented.

Chomsky comments, 'The logic is rather similar to what I have outlined in the case of the acquisition of knowledge of language... small changes in parameters left open in the general schematism can lead to what appear to be very different systems' (1980: 67). To paraphrase Jacob, we could say that what accounts for the difference between English and French, between German and Irish, between Mandarin and Japanese, Kuikúro and Mohawk, Makhuwa and Pitjantjatjara, is

not their structural components; linguistic diversification calls only for different utilization of the same structural information. To put it another way, the parametric genotype determines the surface linguistic phenotype; and, to quote Chomsky (1980) once again: ‘We may think of universal grammar as, in effect, the genetic program, the schematism that permits the range of realizations that are the possible human languages’ (p. 234).²

As that last quotation indicates, both the principles and parameters are seen as part of the genetic endowment, the innate UG. In the context of a minimalist approach to cross-linguistic variation, which seeks to approach the question ‘from below’, we may have cause to question this assumption. One of the central goals of this work is to argue for the ‘emergentist’ approach to parameters, as briefly outlined in the Introduction. The precise nature of that approach will be elucidated in §1.3 to §1.5.

1.1.2 The explanatory value of the P&P approach

Chomsky (1964: 29) made the distinction between observational, descriptive, and explanatory adequacy. Regarding observational adequacy, he says:

Suppose that the sentences

- (i) John is easy to please.
- (ii) John is eager to please.

are observed and accepted as well-formed. A grammar that achieves only the level of observational adequacy would . . . merely note this fact one way or another (e.g. by setting up appropriate lists). (Chomsky 1964: 34)

But, of course, the natural question to ask at this point is why these sentences are different in the way we can observe. To answer this question, we need to move to the level of descriptive adequacy. On this, Chomsky (p. 34) says: ‘To achieve the level of descriptive adequacy, however, a grammar would have to assign structural descriptions indicating that in [(i)] *John* is the direct object of *please* . . . while in [(ii)] it is the logical subject of *please*.’ In fact, we could assign the two examples the structural descriptions in (1):

- (1) a. John_i is $[_{AP} \text{easy } [_{CP} \text{Op}_i [_{TP} \text{PRO}_{arb} \text{to please } t_i]]]$
- b. John_i is $[_{AP} \text{eager } [_{CP} [_{TP} \text{PRO}_i \text{to please } pro_{arb}]]]$

These representations capture the basic difference between the two sentences alluded to in the above quotation, as well as a number of other facts (e.g. that the subject of the infinitive in (1a) is arbitrary in reference; the notional object of *please* in (1b) is arbitrary in reference, etc.).

² It is worth pointing out that the chapter of Chomsky (1980) from which this quotation comes (ch. 5, ‘Language and Unconscious Knowledge’) is based on a talk given on 19 Nov. 1976, and was originally published in 1978.

But linguistic theory must also tell us why these structural descriptions are the way they are. For example, the notations used in (1), CP, PRO, AP, etc., must be explicated. This brings us to explanatory adequacy: we have to explain how the structural descriptions are determined by UG, taking UG to be a theory of possible grammars. If we can do this, we also explain how, in principle, the grammatical properties indicated by representations like those in (1) are acquired (and hence how we, as competent adult native speakers, have the intuitions we have).

As we have seen, the P&P approach allows us to say that a given language is an instantiation of UG with parameters fixed. Many properties of the English *easy-to-please* construction can ultimately be explained in these terms. For example, among the parameters relevant to determining the representation of this construction in (1a) are the following: CP follows the head A (rather than preceding it, as in a head-final language); English has infinitives, and indeed infinitives of this type; arbitrary null pronouns can appear in this context with the properties that we observe them to have; the trace is a *wh*-trace (in many languages, including all the Romance languages, this construction features an A-dependency), etc.

More generally, the P&P approach offered, really for the first time, a plausible framework in which to capture the similarities and differences among languages within a rigorous formal theory. As such, it marked an important step forward in the history of generative grammatical studies. Most important of all, it offered an explanatory model for the empirical analyses which opened a way to meet the challenge of 'Plato's Problem' posed by children's effortless—yet completely successful—acquisition of their grammars under the conditions of the poverty of the stimulus. This becomes particularly clear if we take the view that parametric variation exhausts the possible morphosyntactic variation among languages and further assume that there is a finite set of binary parameters. Imposing an arbitrary order on the parameters, a given language's set of parameter settings can then be reduced to a series of 0s and 1s, i.e. a binary number n . Concomitantly, the task of the language acquirer is to extrapolate n from the PLD. Abstractly, then, the learner can be seen as a function from a set of language tokens (a text, in the sense of formal learning theory stemming from Gold 1967) to n . In this way, we have the conceptual framework for a solution to Plato's Problem and thereby the attainment of explanatory adequacy.

By exactly the same token, the P&P approach offers the promise of a deep language typology. As we mentioned in the previous section, Chomsky (1981: 6) says:

there are certain complexes of properties typical of particular types of language: such collections of properties should be explained in terms of the choice of parameters . . . Ideally, we hope to find that complexes of properties differentiating otherwise similar languages are due to the effects of a single parameter, fixed in one or another way.

Analogously, it offers a framework for research on syntactic change, in that a change in the value of a given parameter may have proliferating effects through the system.

As we have seen, Newmeyer (2005a; 2005b) concluded that the goal of setting up parameter-based typologies was never attained and that P&P had consequently failed. Although Newmeyer raises a number of valid and important criticisms, I consider his conclusion to be premature. I hope to show in what follows that P&P theory should not be abandoned in its general outline; the benefits it offers to linguistic theory are too great for that. But it does need to be substantially rethought and revised, and that is what I set out to do in what follows.³

The adoption of the goals of the minimalist programme requires us to rethink the nature of P&P in any case. First, what we might term ‘methodological minimalism’ requires us to reduce our theoretical machinery as much as possible. This has entailed the elimination of many UG principles, and hence the traditional area of parametrization (Boeckx’s 2015 ‘Parameter’ with a capital ‘P’). We will revisit this point, and its many consequences, in what follows.

Second, there is the matter of ‘substantive minimalism’. This goes beyond merely a severe application of Occam’s razor—methodological minimalism—by asking the question why UG, with the properties we think it has, is that way and not some other way. We can, of course, ask the same question about parameters. Here, the third-factor-driven emergentist approach again offers an attractive answer. Parametric variation is the way it is because of the way the three factors—underspecified UG, PLD, and the acquirer’s characteristic mode of organizing it—interact. The second and third factors conspire to organize variation in a particular way, while the ‘vocabulary’ of variation (e.g. how verb-second is defined over verbs and second positions) is determined by UG. There is an obvious sense in which this takes us beyond explanatory adequacy to a potentially deeper level of understanding. Showing this, by developing a comprehensive account of many aspects of morphosyntactic variation, is the overriding goal of this work.

1.1.3 *The scope of P&P theory*

As we have already mentioned several times, and as was stated very clearly in Chomsky (1981), P&P theory essentially addresses three empirical questions at once. First, it provides a solution to Plato’s Problem, the logical problem of language acquisition, in that the otherwise formidable task of language acquisition is reduced to a matter of parameter setting. Second, it makes predictions about language typology, in that parameters should make predictions about possible language types. Third, it sets the agenda for research on syntactic change, which

³ Chomsky (1980: 142) anticipated this. After commenting on the then-recent inception of the P&P approach, he says:

Each such approach is far too restricted in its basic principles and deductive structure and inadequate in its choice of parameters, and thus remains incapable of accounting for either the richness or the specificity of actual grammars, on the one hand, or their diversity, on the other. Furthermore, each is faced with a vast range of unexplained phenomena and what appears to be—and may actually be—counter-evidence, exactly as we should anticipate for a long time to come, in fact indefinitely, if the field remains a living one.

To make this point more concrete, let us consider a particular domain of variation. One of the best-studied parameters over the history of the P&P approach has been the null-subject parameter (NSP henceforth). We will look at this phenomenon in some depth in Chapter 3, and summarize the history of research on this parameter briefly in §3.1.1 (see also Roberts & Holmberg 2010: 14ff.).

- (2) a. the possibility of a silent, referential, definite subject of finite clauses;
b. 'free subject inversion';
c. the apparent absence of complementizer-trace effects;
d. rich agreement inflection on finite verbs.

(3) Hanno telefonato molti studenti. [Italian]
Have._{3Pl}.Pres telephone.ppt many.MPl students.MPl
'Many students have telephoned.'

(4) *Who did you say that – wrote this book?

(5) Who did you say – wrote this book?

Perlmutter observed that complementizer-trace effects of this kind are not found in null-subject languages (although Rizzi 1982: 121–7 argued that in fact this is not true if certain structures covertly derived at the level of Logical Form are taken

[illegible]

Roberts & Holmberg (2010) further point out that neither Gilligan nor Newmeyer took into account the different kinds of null-subject and null-argument systems that exist; the properties in (2) hold only of consistent null-subject languages of the Italian type, and so many of the 100 languages investigated by Gilligan, e.g. Mandarin Chinese, are not of the relevant type; we will return to this point in §3.2.1. (See Barbosa, to appear, and §3.2 for more detailed discussion of the different types of null-subject languages). For present purposes, however, whatever the merits and demerits of this particular case, the point is that parametric clusters can make typological predictions. Like any typological prediction, these can be falsified by the data.

⁴ In French, the unmarked complementizer introducing finite clauses is *que*, and this obligatorily changes to *qui* when the subject is extracted. See Rizzi & Shlonsky (2007: 130ff.) for an analysis.

beginning with the pioneering study in Hyams (1986). Hyams proposed that early null subjects produced by children acquiring English were an indication of 'parameter mis-setting' in relation to the null-subject parameter, in that children acquiring non-null-subject languages initially set the parameter to the positive value. This led to the suggestion that the null-subject parameter may have a default positive value (while Berwick 1985, motivated by the Subset Principle, argued that the default value was the negative value).

Early null subjects are illustrated in (6) (examples from Guasti 2002: 151; sources are given there):

- (7) a. Se, blomster har. [Child Swedish]
 See, flowers have/has
 'Look, (I/you/she/we) have/has flowers.'
- b. Tickles me.
- c. Mange du pain. [Child French]
 eat-3Sg.Pres some bread

Early null subjects differ from the similar child-language phenomenon of root infinitives in that the verb is clearly finite, as can be seen from the forms in (7), and the fact that they are compatible with the presence of auxiliaries.

Since Hyams' early work, evidence has emerged that early null subjects are not the result of a 'mis-setting' of the null-subject parameter to the 'Italian' value. The main reason for this is that it has been shown that early null subjects do not occur in a range of environments (questions with a fronted *wh*-element, subordinate clauses and matrix clauses with a fronted non-subject) in which null subjects readily occur in adult null-subject languages (see the discussion in Guasti 2002: 159–60; 2016: 211–12).

Another option, pursued by Hyams (1992), was to claim that early null subjects result not from a 'subject-drop' option of the familiar Italian kind, but from a 'topic-drop' option of the kind seen in languages such as Chinese and Japanese (see in particular Huang 1984; 1989; Huang & Roberts 2017: 313–14). The advantage of this idea is that it reconciles the occurrence of early null subjects in languages with impoverished agreement systems with the known facts of adult languages: while null-subject languages like Italian appear to require 'rich' verbal agreement for the recovery of the content of null subjects, topic-drop languages like Chinese and Japanese have no agreement at all and yet allow null arguments of various kinds; see §3.2.4. The disadvantage of this approach is that topic-drop languages typically allow null objects fairly freely, while early null objects are rather rare in child English (Hyams & Wexler 1993); for further discussion and statistical evidence, see Guasti (2002: 157–8; 2016: 212). So the idea that the putative parameter mis-setting is in the 'Chinese' direction rather than in the 'Italian' one does not appear to hold up either.

Other possibilities which have been explored to account for this phenomenon include relating it to the 'diary drop' phenomenon discussed in particular by Haegeman (2000; 2007) in relation to English and French. This is illustrated in (8):

- (8) a. Cried yesterday morning. (Plath 1983: 28)
 b. Elle est alsacienne. Paraît intelligente. [French]
 She be.3Sg.Pres Alsatian.FSg. Seems.3Sg.Pres intelligent.FSg
 ‘She is from Alsace. Seems intelligent.’ (Léautaud 1989: 48)

These examples have been argued to involve clausal truncation by Haegeman (2000) and Rizzi (2000); see Guasti (2002: 166ff.; 2016: 215–17) for summary and discussion. The other accounts of early null subjects which have been put forward involve extra-syntactic factors, such as processing difficulties (Bloom 1990) and metrical difficulties, i.e. dropping of weakly stressed syllables, which subject pronouns typically are (Gerken 1991). Again, the purpose here is not to evaluate these proposals, but to observe that the proposed NSP has led to productive work in language acquisition.

Finally, the NSP has also stimulated proposals regarding syntactic change. This is particularly clear in the case of the history of French. At earlier stages of its history, French was a null-subject language. The value of this parameter seems to have changed around 1600. In Old French (OF, 842–1300) and Middle French (MidF, 1300–1500) we can readily find examples of null subjects, such as the following (from Roberts 1993a: 124f.):

- (9) Old French:
 a. Tresqu’en la mer cunquist la tere altaigne.
 Until the sea conquer.3Sg.Pst the land high
 ‘He conquered the high land all the way to the sea.’ (Roland, 3)
 b. Si chaï en grant povreté.
 Thus fall.1sg.Pst into great poverty
 ‘Thus I fell into great poverty.’ (Perceval, 441)
- (10) Middle French:
 a. Et ly direz que je me recommande
 and her.Dat say.2Pl.Fut that I myself recommend.1Sg.Pres
 humblement a elle
 humbly to her
 ‘And you will say to her that I humbly ask her good will.’ (S 131, 16)
 b. Ne vous pourroye a demi dire le tresgrant dueil
 neg you can.1Sg.Cond at half say the very-great grief
 ‘I could not tell you half the great grieving.’ (S 165, 17; Vance 1997: 260)

The pioneering work in this area was Adams (1987a; 1987b). Later work by Roberts (1993a) and Vance (1989; 1997) developed these ideas further. One important issue is the connection between the NSP and the fact that OF was verb-second; this was already noticed in the traditional historical-linguistic literature (Thurneysen 1892), and was developed by the above authors, as well as by Kaiser (2002), Sitaridou (2012), and, most recently in a wider Romance perspective, Wolfe (2015; 2019). There has also been important recent work on null

subjects in older Germanic languages (see Walkden 2014: 157ff.; Kinn 2014; 2016; Kinn, Rusten, & Walkden 2016; §3.2.3, and the references given there).

So we see that the postulation of the NSP has led to new typological investigations, new discoveries in child language, and a new approach to studying historical change. A brief survey of this kind could be repeated for other parameters proposed in the GB era; the head parameter, the *wh*-movement parameter, various verb-movement parameters, and others (see Roberts 2007a: ch. 1 for a survey of proposed changes affecting a range of parameters, and Huang & Roberts 2017: 308–18 for discussion of several GB-era parameters). Here we see the contribution the P&P approach has made to linguistic theory. At the very least, it has proven to be an excellent heuristic, but it clearly offers so much more. This is what motivates the present work; the notion of parameter of UG must be recast so as to be compatible with minimalist approaches and so as to meet some of the criticisms that have been made.

1.2 Questions for classical parameter theory: the Romance languages

In this section, I will illustrate the task before us by looking in some detail at aspects of syntactic variation in the Romance languages. The Romance languages could be thought of as the ‘home of the microparameter’, in that it is largely in relation to these languages that that approach has developed; we will discuss microparametric approaches in detail in §1.3.1. Kayne (2005a) argues for a microparametric approach on the grounds that, in order to isolate individual parameters of variation, we need to hold as many potentially variable properties (i.e. other parameters) as possible constant. Since we know so little for certain about parametric interactions, the more different two languages are, the more difficult it will be to isolate a single point of variation. In an ideally controlled experiment in parametric variation, then, we would want to isolate just one parameter and look at its effects (note that this is relevant to Newmeyer’s critique of the NSP cluster mentioned in the previous section; the critique assumes, without real grounds, that other variables are controlled; see Roberts & Holmberg 2010: 19–20 for discussion and elaboration of this point). As Kayne says, ‘the probability of discovering what syntactic property is parametrically linked to what other one... is higher when the “experiment” approaches to a greater extent the ideal of the controlled one’ (2005a: 8). Looking at a fairly closely related set of well-described languages such as the Romance family, therefore, fits the bill for the kind of microparametric investigation Kayne advocates.

There is nothing special (synchronically) about the Romance languages as a family that leads me to use them as an example here. They simply happen to be a well-documented group of fairly closely related languages. For various contingent reasons, they have been very intensively studied from the P&P perspective. Works on Romance syntax from a P&P perspective are too numerous to list in anything like an exhaustive way, but it is important to mention the pioneering works of

Kayne (1975), Burzio (1986), Jaeggli (1982), and Rizzi (1982); indeed these, along with Chomsky (1981), launched the entire P&P enterprise. It is also worth mentioning the monumental three-volume survey of Italo-Romance dialects by Manzini & Savoia (2005), and the comprehensive collection edited by Ledgeway & Maiden (2016).

Of course, the Romance languages do offer us something special from the diachronic perspective, as they represent one of the very few language families whose (almost) direct parent language is well attested, in the form of Latin. It is known that Classical Latin was a slightly artificial, probably somewhat archaizing, literary language (see Clackson 2011; Clackson & Horrocks 2007: 183ff.; Ledgeway 2012: 202f.), but the corpus of texts is large and there are very valuable texts in various kinds of Late Latin, which may more directly reflect the nature of the common parent(s) of Modern Romance (Adams 2007; 2011; Clackson 2004). Moreover, the Romance languages are fairly well documented throughout their history (wherever one chooses to draw the ultimately arbitrary line between Late Latin and Early Romance; see Wright 2011: 64 for discussion), a further invaluable resource for diachronic syntax, as the brief comments on the history of French in the previous section indicate.

So in this section I want to illustrate in detail the kind of variation that we can find in the Modern Romance languages in relation to a range of salient syntactic phenomena. What we will see is very intricate micro-variation necessitating the postulation of numerous (micro)parameters. The discussion goes into some detail (although much more can be found in the sources cited), in order to show what the nature of the true micro-variation that has been uncovered is. The reason for going into this material in this much detail is not to describe or analyse the Romance phenomena in themselves (the sources cited do this), but rather to illustrate several important points about morphosyntactic variation that any theory of that variation must account for. These are, first, that since there is nothing intrinsically special about these languages, they serve as an illustration of the kind of variation that we can find anywhere (China, Tanzania, etc.). Second, there is also much that is *not* variable in Romance, but what we know can vary in the world's languages. This last point will be illustrated by a comparison of Romance with Japanese. Third, there are the properties which do not vary either in Japanese or in Romance: the question then becomes that of developing a theory of parametric variation which can account for both the similarities and the differences among the Romance languages, as well as the differences from Japanese (and, of course, the similarities between Romance and Japanese). This discussion sets the stage for the discussion of different parameter types in §§1.3 and 1.4.

1.2.1 *Variation in subject clitics*

The variation in subject clitics (SCLs) across Romance is quite significant. Indeed, it is difficult to provide a straightforward definition of SCLs. Poletto & Tortora (2016: 773) define them as follows:

the term ‘subject clitic’ encompasses an assortment of clitic morphemes instantiating a disparate array of functional heads which have in common two things: (a) a history of being grouped together as a class of elements (correctly or not), and (b) their appearance in what we could term ‘the higher functional field’, that is, the portion of the clause in which inflectional (‘higher IP’...) and information-structural (‘lower CP’...) information is expressed.

They also point out that subject clitics are largely geographically limited to ‘the northern half of Italy, southern France, and Switzerland. Thus, subject clitics are claimed to be found in northern and southern Gallo-Romance varieties, Franco-Provençal, Rhaeto-Romance, and the dialects of northern Italy and northern Tuscany.’

The class of SCLs is somewhat heterogeneous, as just pointed out. A first distinction can be made between (Standard) French and the general Northern Italian situation. This can be seen if we compare the following paradigms:

- (11) i. a dǝrmə, tə dǝrmə, i/al dǝrmə, a durmiŋ, durmitə, i/al dǝrmənə
(Carrara, N. Italy; Manzini & Savoia 2005: I.80).
- ii. dǝrmi, ta dǝrmat, al/la dǝrma, dǝrmum, dǝrmuf, dǝrman
(Como, N. Italy; Manzini & Savoia 2005: I.100).
- iii. je dors, tu dors, il/elle dort, nous dormons, vous dormez, ils/elles
dorment (French).
All: ‘I sleep, you sleep, he/she sleeps, etc.’
- iv. (E) parlo, tu parli, e/la parla, si parla, vu parlate, e/la parlano
‘I/you/(s)he/we/you/they speak/speaks.’
(Fiorentino; Brandi & Cordin 1989).

Here we observe that only French has a full paradigm of subject clitics, present and distinct in all person and number combinations. Since Kayne (1975; 1983), it has been thought that French subject clitics are full syntactic phrases, presumably DPs (although Cardinaletti & Starke 1999 analyse them as ‘weak pronouns’ and as such structurally deficient as compared to typical definite non-pronominal DPs). In this, they do not differ significantly from non-clitic pronouns in a language like English (which also, in many cases, have weak and strong versions, as Cardinaletti & Starke show). The typical Northern Italian SCLs, on the other hand, are usually thought to be syntactic heads. One property that supports this distinction can be readily observed in (11): while the French SCLs are distinct for each person–number combination, as one might expect of a system of personal pronouns, none of the Northern Italian ones shows such a six-way distinction in its SCLs. Moreover, each of the Northern Italian varieties behaves slightly differently, in that Carrarese has SCLs in all persons except 2Pl, although it displays two syncretisms (1Sg/1Pl and 3Sg/3Pl), while the Como variety only has SCLs in 2Sg and 3Sg and Fiorentino lacks an obligatory SCL only in 1Sg.

Roberts (2010a:106) introduces the feature $[\pm\text{agr}]$ as a descriptive device to denote whether a clitic or agreement paradigm shows a full set of morphological

person–number distinctions (a ‘full’ set of distinctions may contain at most one zero exponent and one syncretism, which may be the zero exponent). He then makes two further assumptions: (i) that a ‘pronominal’ paradigm must be a full paradigm in the sense just defined, and (ii) if verbal inflection shows a ‘pronominal’ paradigm, then the system is a null-subject one. Using the $[\pm\text{agr}]$ notation, we then have the following set of possibilities:

- (12) a. SCL[+agr] V[+agr] a ‘fully redundant,’ null-subject system
- b. SCL[+agr] V[–agr] a non-null-subject system
- c. SCL[–agr] V[+agr] a non-redundant null-subject system
- d. SCL[–agr] V[–agr] (usually) a complementary system

Fiorentino instantiates a (12a)-type system, in which the SCLs and the verbal inflections covary, with only a small amount of syncretism in the clitic paradigm (involving 1Sg/3MSg/3MPI *e*) and one (complementary) syncretism in the verbal paradigm (between 3Sg and 1Pl). Here, then, the SCLs are a realization of the unvalued ϕ -features of T, just like the verbal inflection, and the system has the ‘rich’ agreement characteristic of null-subject languages. The single difference with a language like Standard Italian or Spanish is that the agreement is doubly marked, once by the SCLs and once by verbal inflection.

French is an example of (12b). If the French SCLs are regular pronouns, then this kind of system is a non-null-subject system.⁵ Roberts (2010a: 107) points out that just one of the 180 Italian dialects reported by Manzini & Savoia (2005: I.371ff.) shows this pattern (with the pronoun in proclisis, and leaving aside the verb-second Rhaeto-Romance varieties).

Dialects of the Como type, with fully differentiated verbal inflection but syncretisms and gaps in the clitic paradigm, represent type (12c). Finally, a fairly common pattern is that schematized as (12d), where neither the SCL paradigm nor the verbal-inflection paradigm alone shows a full set of forms, but together they form a single complementary (or near-complementary) pattern (as was already observed by Renzi & Vanelli 1983 and Poletto 2000). This is exemplified by the Carrarese paradigm in (11i). Roberts (2010a:108) argues that these are also null-subject languages, with distributed ‘rich agreement’ (see in particular Roberts 2014a: 195f. for an analysis using mechanisms from Distributed Morphology).

Poletto & Tortora divide the SCLs of all varieties other than Standard French into four subclasses, following Poletto (2000). The classification leaves aside a number of further distinctions (see in particular Manzini & Savoia 2005: I.69f., 122–62) and is limited to SCLs found in declarative clauses with simple tenses: interrogative and auxiliary SCLs show further properties. The four classes are person clitics, number clitics, ‘deictic’ clitics, and invariable clitics.

Person clitics are restricted to 2Sg and 3Sg (defined, Poletto & Tortora suggest, by the feature [Hearer]). They typically have the form *tV* and *Vl*, where *V* is any vowel. They have the following properties:

⁵ Non-Standard varieties of French may well be more similar to the Northern Italian varieties (Roberge 1986; Roberge & Vinet 1989; Zribi-Hertz 1994; Sportiche 1999; Culbertson 2010; Culbertson & Legendre 2008; Roberts 2010b).

(a) they always occur to the right of the preverbal negative marker; (b) they must be repeated in all types of coordination...; (c) they never interact with left-peripheral elements (such as the complementizer);[footnote omitted] (d) they are never found in enclisis...; and (e) they occur to the right of all other types of subject clitics.

(Poletto & Tortora 2016: 775)

The 2Sg *tu* clitic in Fiorentino is an example of this type of SCL in that it is preceded by the clausal negator *un* (*Un tu mangi* ‘You do not eat’; see §7.3.2 for more on the position of clausal negators in Romance and elsewhere).

Number clitics are typically restricted to 3rd person (singular and plural) and often show gender distinctions. They differ from person clitics in that ‘they can be found either to the right or to the left of the negative marker’ and they must be repeated in certain types of coordination, but not others. Furthermore:

(c) like person clitics, they do not interact with left-peripheral elements, but unlike person clitics, they do cluster with the complementizer in embedded clauses; (d) in contrast with person clitics, they are found in enclisis in interrogatives;[footnote omitted] (e) if they co-occur with invariable or deictic subject clitics... they always occur to their right.

(Poletto & Tortora 2016: 776)

Deictic clitics occur in all persons, but typically with syncretisms across 1st and 2nd person and between singular and plural in all persons. They are usually vocalic in form. The Emilian variety of Gainago/Torrile (province of Parma; Maria Pedretti, p.c.) illustrates:

- (13) (a) *dormi, a t dormi, a l/la dorma, a dormome, a dormiv, i dormen.*
‘I sleep’, ‘you sleep’, ‘he/she sleeps’, etc. [Gainaghese]

(Note the presence of the person clitics *t, l/la* in the 2Sg and 3Sg, and the complementary syncretisms of the verbal inflection in 1Sg and 2Sg and the clitics in 1Sg, 1Pl, and 2Pl.) Deictic clitics (a) ‘always occur to the left of the preverbal negative marker’; (b) can be omitted in certain kinds of coordination (but different kinds as compared to number clitics); ‘(c) unlike person and number clitics, interact with left-peripheral elements;[footnote omitted] but unlike person clitics, do cluster with the complementizer in embedded clauses; (d) in contrast with number clitics, are never found in enclisis in interrogatives; (e) like invariable clitics..., always appear to the left of person and number clitics’ (Poletto & Tortora 2016: 777).

Finally, invariable clitics do not, as the name implies, overtly encode any person or number features. They may appear with all persons, but do not do so in all varieties. They have the following properties:

(a) like deictic clitics, they always occur to the left of the preverbal negative marker; (b) they can be omitted in all types of coordination; (c) unlike person and number clitics (but like deictic clitics), they interact with left-peripheral elements;[footnote omitted] (d) like deictic clitics, they are never found in enclisis in interrogatives or in any other construction where subject clitic inversion is known to occur; (e) like deictic clitics, they always appear to the left of person and number clitics. (Poletto & Tortora 2016: 778)

The variation in the syntactic contexts in which SCLs of the various kinds can appear is considerable. No variety has them in imperatives or non-finite contexts (the latter fact, at least, supporting the idea that they are in part agreement elements). Some varieties show ‘subject–clitic inversion’ in interrogatives, conditionals, and exclamatives (see also Munaro 2010 and §§5.6, 5.7 on residual V2), such as the Friulian (of Clauzetto) and Paduan varieties in (14) (Poletto & Tortora’s 2016: 780 (23b,c)):

- (14) a. Vinisi-al tjo pari, o
Come.3Sg.Cond=SCL.3Sg your father SCL.1Pl
podaresin là. [Friulian]
can.1Pl.Cond go.Inf
'If your father came, we could leave.'
- b. Quanto belo se-lo! [Paduan]
how nice be.3Sg.Pres=SCL.3Sg
'How nice it is!'

(15) concessive > counterfactual > exclamative > interrogative

The enclitics that appear in these contexts often show different forms from the proclitics seen in declaratives and, in many cases, show a fuller range of forms (an observation also made by Renzi & Vanelli 1983). While inversion is being (or has been) lost in many varieties, and the same can be observed for varieties of French and indeed English (see Biberauer & Roberts 2017b and §5.6 on conditional inversion in the history of English), there are Lombard varieties where apparent inversion has generalized to declarative clauses (Poletto & Tortora 2016: 780, (24c)):

- A further case of this is the phenomenon of 'OCL-for-SCL' found in some Piedmontese and Franco-Provençal varieties. In these varieties the SCL appears on the auxiliary unless a proclitic OCL does, in which case the SCL disappears.

(19) illustrates this from the Franco-Provençal variety of Ayas in the Val d'Aoste (from Roberts 2016: 792):

- (19) a. Gnunc l' a viu -me. [Franco-Provençal, Ayas]
 noone SCL.3Sg=have.3Sg.Pres see.ppt=OCL.1Sg
 b. Gnunc m' a viu.
 no one OCL.1Sg=have.3Sg.Pres see.ppt
 'No one has seen me.'

Last of all, some SCLs have left-peripheral functions: Paduan *a* marks new information (Benincà 1983), Donceto *-v* marks a clause as interrogative (Cardinaletti & Repetti 2008), and some Friulian SCLs only appear with tonic *wh*-phrases (Poletto 2000).

It is likely that, if anything, the above summaries do not do justice to the attested variation in SCLs in this part of the Romance-speaking area (see also Manzini & Savoia 2005: I.122ff.). Furthermore, there is no doubt further variation which has not yet been documented. The variation observed above concerns, first, the presence of SCLs (French, Franco-Provençal, some varieties of Occitan, Northern Italian, Northern Tuscan, and some Rhaeto-Romance: yes; the rest of Romance: no); then, whether the SCLs are heads (not Standard French, but all the other varieties just mentioned with the possible exception of Rhaeto-Romance); then there appears to be a cross-cutting set of options concerning the feature content and realization of a range of functional heads in the inflectional field and the left periphery. These appear to involve at least four heads (one to host each of the basic SCL types identified by Poletto 2000 and Poletto & Tortora 2016), and a range of features involved (or not) with licensing different kinds of subjects in different contexts: these clearly include φ -features (with different bundles being associated with different heads), *wh*-features, and features sensitive to definiteness and/or quantification. The implicational scale in (17) represents five options, and there is the further one of whether SCLs license null subjects at all. We mentioned three different left-peripheral licensing options and two different 'placeholder' options (see also the discussion of the different kinds of OCL-for-SCL in Roberts 2016: 792). According to Munaro's (2010) implicational hierarchy for SCL inversion given in (15), there are five options in this domain (counting the 'no-inversion' option): these involve clause-typing features of the left periphery such as Q, Realis features, and exclamative features. All in all, then, at least the distribution and realization of formal features for Person, Number, Gender, Deixis, various clause-types (including new information, Q, Realis and Exclamative), *Wh*, and Definiteness are implicated in determining the incidence and distribution of SCLs. It is hard to estimate the range of overall variation available, but it is clear that there is a non-trivial number of parametric options.

This is not the place to present a detailed analysis of the variation we observe in Romance SCLs. To do that would be a book-length project in itself, as should be apparent from the above description. The purpose of the discussion in this section is to illustrate the range of variation in SCLs that has been described and, to varying degrees, analysed in the literature. Since SCLs clearly interact with null

subjects, it has long been thought—since the pioneering work of Renzi & Vanelli (1983) and Rizzi (1986a)—that this variation represents parameter settings, either a subcase of the null-subject parameter or a closely related parameter. It is abundantly clear that there is rich and intricate variation in these languages and varieties in this domain.

1.2.2 Variation in negation

The parameters regulating aspects of negative concord and clausal negation will be discussed in §7.3; here I limit the discussion to the illustration of variation in negation in Romance. Across the Modern Romance languages, we find all three options for the expression of clausal negation allowed by ‘Jespersen’s Cycle’: a preverbal negator only, usually a clitic, a combination of preverbal clitic and postverbal ‘adverbial’ negation, and just postverbal negation. The following examples illustrate (see Zanuttini 1997: 14):

- | | | | | | |
|------|----|-----------------------|----------------|----------------|---------------------|
| (20) | a. | Non | dormo. | | [Standard Italian] |
| | | NEG | sleep.1Sg.Pres | | |
| | b. | Je | ne | dors | pas. |
| | | I | NEG | sleep.1Sg.Pres | NEG |
| | c. | i | | drøma | mia |
| | | SCL.1Sg | sleep.1Sg.Pres | NEG | |
| | | All: ‘I don’t sleep.’ | | | |
| | | | | | [Trecate, N. Italy] |

The first option is found in Standard Italian, all of Ibero-Romance, Romanian, and Central and Southern Italo-Romance (thanks to Adam Ledgeway, p.c., for information on this last group of varieties). The second is best known from Standard French, but is also found in many Northern Italian dialects, for example those of Lombardy and the Po Valley (Manzini & Savoia 2005: III.133ff.). The third is found in many Northern Italian dialects and elsewhere (see below). It has been known since Jespersen’s pioneering work that there is a natural diachronic tendency for systems to develop from (20a) to (20c) via (20b) (for detailed discussion of Jespersen’s Cycle in a range of languages, see Willis, Lucas, & Breitbarth 2013: 9–13). So we can immediately observe variation in the position (pre- vs postverbal) and the expression (one negative morpheme or two) of negation across Romance. These two kinds of variation—position and expression—are the two principal dimensions of parametric variation associated with negation (see §7.3.2), the second being connected to the phenomenon of negative concord, on which, see below (see also §7.3.1 on the parameters regulating negative concord).

As usual, a closer look reveals that each dimension of variation is more complex, revealing subtypes of various kinds across Romance. Poletto (2016) documents this in some detail, and the following exposition is based on hers, although a number of details are omitted for simplicity here.

Concerning preverbal negation, there is variation across the Northern Italian dialects concerning the position of this element in relation to SCLs, as we saw in

(21) I mituns ne no vègn nia
The boys not not come.3Pl.Pres not
encò. [Rhaeto-Romance, San Vigilio di Marebbe]
today
'The boys are not coming today.'
(Poletto 2016: 835)

(22) Non piove (mica). [Northern Regional Italian]
Not rain.3Sg.Pres not
'It's not raining.'

Zanuttini (1997) shows that, on closer observation, there are three distinct positions for postverbal negators (see §7.3.2, Table 7.1, for further details and for more on the positions of clausal negators across languages). A first position, characteristic of French *pas*, is before that occupied by adverbs of the 'already' class, as identified by Cinque (1999: 4–11). A second one follows these adverbs but precedes the position for the 'no longer/always' class; this is characteristic of Piedmontese *nen*, for example. Finally, a third position follows all adverbs; this characterizes Brazilian Portuguese *não*. There appears to be an interesting partial correlation between the etymologies of the negators and which of these positions they occupy. The first class are typically derived from minimizers,

words denoting very small quantities (*mica* comes from ‘crumb’, *pas* from ‘step’, etc.). The second are etymologically negative quantifiers and/or n-words (Piedmontese *nen* derives from the word for ‘nothing’, *niente* in Standard Italian). The third class is identical in form with the basic preverbal sentential negator (see Zanuttini 2010; Poletto 2016: 839; Manzini & Savoia 2012 for evidence that the correlation is not exact). Furthermore, there are some Northern Italian varieties, e.g. Veneziano, in which the second type of postverbal negator is sensitive to the aspectual type of the verb, appearing with activities but not accomplishments (Poletto 2016: 842):

- (23) a. No la salta gnente, sta ranetta de
 Not SCL.F3Sg jump.3Sg.Pres nothing, this frog of
 carta. [Veneziano]
 paper
 ‘This paper frog does not jump at all.’
- b. *No la salta zo gnente, sta ranetta
 Not SCL.F3Sg jump.3Sg.Pres down nothing, this frog
 de carta.
 of paper

In addition to interacting with aspect, negation also interacts with modality. In certain Southern Calabrian dialects, negation interacts with the complementizer-like sentential particle *mu* which introduces a subjunctive clause (see Roberts & Roussou 2003: 88f., who argue that these particles are very similar to Greek *na* and as such are in a left-peripheral position corresponding to Rizzi’s 1997 Fin). Negation combines with *mu* to form the complex *nommu*, as in (Damonte 2008, cited in Poletto 2016: 842):

- (24) Eu speru nommu lejunu a to
 I hope.1Sg.Pres not.Subjunc read.3Pl.Pres the your
 littera. [Locri, Calabria]
 letter
 ‘I hope they will not read your letter.’

Finally, Poletto (2016: 843) points out that certain adverbs seem to ‘agree’ with negation in negative clauses: Veneto *ancora* (‘still, yet’) becomes *gnancora*, and Piedmontese *nen* combines with *pi(ü)* (‘any longer’), giving variously *pinin*, *pinen*, or *piügn*.

Turning to negative concord, clearly in languages of the (20b) type negative form does not map directly to negative semantics; one of the two negators must fail to carry semantic negation (this is usually thought to be the preverbal one, as it can appear alone in certain contexts without negative semantics, e.g. in comparatives; see §7.3.1). The Romance languages also show both strict and non-strict negative concord, Romanian being of the former type and Italian, Spanish, and Portuguese of the latter (Catalan seems to vary; see again §7.3.1):

- (25) a. Nessuno (*non) ha telefonato. [Italian]
 No one not have.3Sg.Pres phone.ppt
 'No one called.'
- b. Niciun student nu a venit. [Romanian]
 No student not have.3Sg.Pres come.ppt
 'No student has come.'

French is also a strict negative-concord language in regard to *ne*, but not *pas*, which, if added to (26), gives rise to a double-negation reading (Déprez 1999):

- (26) Personne n' est venu. [French]
 No one neg=be.3Sg.Pres come.ppt
 'No one has come.'

Haitian Creole, on the other hand, lacks *ne* but has negative concord, not double negation, with *pa* (Déprez 1999):

- (27) M pa we pèsonn. [Haitian Creole]
 I not see no one
 'I did not see anyone.'

Note that *pa* is preverbal here; this is because Haitian Creole lacks V-to-T movement of the kind found in finite clauses in Standard French (Pollock 1989; DeGraff 1997; 1999a; 1999b; Roberts 1999; §5.3.2). Apparently, then, *pa* here occupies the same position as French *pas* (but see §5.3.2 for evidence that this is not in fact the case). See §7.3.1 for more details on the parameters governing negative concord and the associated parameter hierarchy.

Once again, we see a range of variation involving the position of the clausal negator, its interpretation, and how it interacts with aspect, mood, and adverbs. We also see the two main types of negative concord. While no Modern Romance language lacks negative concord, it is well known that Latin, like Modern Standard English, did lack negative concord (Gianollo 2018). We thus observe three principal patterns of clausal negation (with various subtypes conditioned by aspect and mood in particular) and two types of negative concord.

1.2.3 Variation in enclisis of object clitics

We saw in §1.2.1 that only a subgroup of Romance languages has subject clitics. On the other hand, nearly all the Modern Romance languages have complement clitic pronouns, except for the majority of creoles, and Colloquial Spoken Brazilian Portuguese (Lucchesi & dos Passos-Mendes 2009: 474–6, 480–4). Here I limit the discussion to direct-object clitics. The cross-linguistic incidence of indirect-object clitics is close (but not identical) to that of direct-object clitics; there is considerable variation, however, regarding the genitive and locative clitics (e.g. these clitics are not found in Modern Spanish, Portuguese, or Romanian). Direct-object clitics are special clitics in the sense of Zwicky (1977), since they generally occupy a position distinct from that of a stressed object pronoun or a non-pronominal

direct object. The Romance languages fall into four main types regarding the distribution of enclitics and proclitics (see also Shlonsky 2004).

The first type is represented by European Portuguese, Galician, and Asturian; I illustrate with European Portuguese throughout. Here enclisis is more widespread than elsewhere in Modern Romance. Enclisis is found with a definite subject in positive declarative matrix clauses:

- (28) O Pedro (**a) encontrou -a. [European Portuguese]
 the Pedro OCL.3FSg=meet.3Sg.Past=OCL.3FSg
 'Pedro met her.'

Proclisis is found in embedded clauses, in negated declarative matrix clauses, with an initial negatively quantified constituent, an initial focused constituent, and certain initial quantified constituents:

- (29) a. Dizem que o Pedro a
 Say.3Pl.Pres that the Pedro OCL.3FSg
 encontrou(**-a). [European Portuguese]
 =met =OCL.3FSg
 'They say that Pedro met her.'
- b. O Pedro não a encontrou(**-a).
 the Pedro not OCL.3FSg=meet.3Sg.Pst=OCL.3FSg
 'Pedro didn't meet her.'
- c. Ninguém me ajudou (**-me).
 nobody OCL.1sg=help.3Sg.Pst=OCL.1sg
 'Nobody helped me.'
- d. Onde a encontrou (**-a) o Pedro ?
 where OCL.3FSg=meet.3Sg.Pst=OCL.3FSg the Pedro
 'Where did Pedro meet her?'
- e. Até o Pedro me deu (**-me) uma prenda.
 even the Pedro OCL.1Sg=gave =OCL.1Sg a gift
 'Even Pedro gave me a gift.'
- f. Todos os rapazes me ajudaram(**-me).
 All.MPl the.MPl boy.MPl OCL.1Sg help.3Pl.Pst=OCL.1sg
 'All the boys helped me.'

See Madeira (1992; 1993; 1994), Rouveret (1992), Martins (1994), Uriagereka (1995), Costa (2001), Duarte and Matos (2000), Raposo (2000), Shlonsky (2004), and Raposo & Uriagereka (2005). If, following Kayne (1991), we think of the clitic as occupying a constant position, then the proclisis/enclisis alternations must involve verb movement. Where we have enclisis, the verb has raised to a higher position than where we have proclisis; in fact, we can think of the various pre-clitic elements in (29) (complementizers, negation, *wh*, focus, and certain quantifiers) as either directly blocking the 'high' verb-movement which gives rise to enclisis or as being associated with left-peripheral heads which do this (see Roberts 2012b for an analysis along these lines; this kind of verb-movement is thus a case of residual

verb-second: see §§5.6, 5.7). A consequence of this is that definite subjects, as in (28), are high in the left periphery; Barbosa (1995; 2009; to appear) argues that they are left-dislocated (with the ‘rich’ subject-agreement on the verb acting like a resumptive clitic in clitic left-dislocation; see §3.2.2).

In non-finite clauses, there are complex and intricate patterns of enclisis in European Portuguese depending on whether the infinitive is inflected or not, the nature of the selecting verb (epistemic, factive, etc.), negation, and whether the infinitive is a subject or an adjunct, and, if a subject, whether it is introduced by a preposition (see Madeira 1994; Raposo & Uriagereka 2005; and for an overview, see Roberts 2016). Support for the idea that the verb moves in at least some of these cases comes from subject-inversion (Raposo 1987). There is a fairly consistent pattern of enclisis with subject-inversion; since subject-inversion is standardly analysed, in many languages (including English), as involving verb-movement into the left periphery (an analysis that originates with den Besten 1983; see §§5.5.2, 5.6), then we expect enclisis to correlate with this.

European Portuguese also shows ‘mesoclisis’: what appears to be cases where the clitic appears ‘inside’ a complex verb, between the verb stem and the tense/agreement inflection:

- (30) Escrevê-lo -ei. [European Portuguese]
 write =OCL.3MSg =fut.1Sg
 ‘I will write it.’

An initial focused constituent, as before, leads to obligatory proclisis (Raposo 2000: 284). This phenomenon was found more widely in the medieval Romance varieties. One possible analysis (suggested in Roberts 1993b) is to treat the apparent future-tense marker as an enclitic auxiliary; in that case mesoclisis reduces to enclisis of the pronominal clitic combined with further enclisis of the auxiliary in the given order.

The second type of OCL system found in Romance involves general proclisis in finite clauses, but enclisis in a range of non-finite contexts. In Spanish and Italian, the general order of clitics in relation to clause-mate infinitives, in all types of infinitive contexts, is enclisis:

- (31) a. Gianni vuole mangiarlo. [Italian]
 b. Juan quiere comerlo. [Spanish]
 Gianni/Juan want.3Sg.Pres eat.Inf=OCL.3MSg
 ‘Gianni/Juan wants to eat it.’
 c. Gianni ha deciso di mangiarlo. [Italian]
 d. Juan ha decidido comerlo. [Spanish]
 Gianni/Juan have.3Sg.Pres decide.ppt of eat.Inf=OCL.3MSg
 ‘Gianni/Juan has decided to eat it.’
 e. Parlargli sarebbe un errore. [Italian]
 f. Hablarle sería un error. [Spanish]
 talk.Inf=DatCL.3MSg be.3Sg.Cond a mistake
 ‘To talk to him would be a mistake.’

Here we see enclisis to infinitives in the complement to restructuring and non-restructuring verbs, as well as a non-complement infinitive. The same pattern is found in many Italo-Romance varieties (Manzini & Savoia 2005: III.335–57).

On the other hand, proclisis is the norm in infinitives in French and Brazilian Portuguese, as well as in many Southern Italian dialects, and there is no clitic-climbing (except in French causatives; see Kayne 1975 and §6.4.1). This is the third type of system:

- (32) a. Jean veut le manger. [French]
 Jean want.3Sg.Pres OCL.3MSg=eat.Inf
 ‘John wants to eat it.’
 b. Jean a décidé de le manger. [French]
 Jean have.3Sg.Pres decide.ppt of OCL.3MSg=eat.Inf
 ‘John has decided to eat it.’

Certain Italo-Romance varieties show an interesting alternation of proclisis and enclisis in infinitives. A number of Tuscan, Po Valley, Ligurian, and Sicilian dialects have optional or obligatory proclisis in negative infinitives, and a subset of these have proclisis in *wh*-infinitives. The following examples from Modenese (from Manzini & Savoia 2005: III.357, stress-marking removed) illustrate the interaction with negated infinitives:

- (33) a. a t ɔ det... [Modenese]
 SCL= OCL.2Sg have.1Sg.Pres say.ppt
 ‘I told you ...’
 b. ...d andɛr -əg
 of go.Inf =LocCL
 ‘...to go there.’
 c. ...d an g andɛr (briza)
 to neg Loc.CL=go.Inf (neg)
 ‘...to not go there.’

Following what we said about enclisis in finite contexts above, enclisis to infinitives involves infinitive-movement (this idea is developed at length in Kayne 1991).

Finally, the fourth pattern is found in the Piedmontese variety Borgomanerese (Tortora 2000; 2002; 2010; 2014a; 2014b) and other closely related varieties (Trecate, Galliate, Cerano, Quarna Sotto; see the references and discussion in Tortora 2014b: 83). These varieties have general enclisis to the verb in finite contexts. I illustrate with Borgomanerese (Tortora 2010: 137f.):

- (34) a. La môngia -la. [Borgomanerese]
 SCL.3FSg=eats3Sg.Pres OCL.3FSg
 ‘She eats it.’
 b. I mcengia -la sempri.
 SCL.1Sg=eat.1sg.Pres=OCL.3FSg always
 ‘I always eat it.’

(35) a. I porti mi -lla. [Borgomanerese]
SCL.1Sg=bring.1Sg.Pres neg=OCL.3FSg
'I'm not bringing it.'

b. I voenghi piö -lla.
SCL.1Sg=see.1Sg.Pres more=OCL.3FS
'I don't see her any more.'

c. *I trati mal -lu.
SCL.1Sg=treat.1sg.Pres badly=OCL.3MSg

(36) Jean l' a vue. [French]
 John OCL.3FSg=have.3Sg.Pres see.ppt.FSg
 'John has seen her.'

(37) Dz' i batia -la tot solet. [Valdôtain]
 SCL.1sg=have.1Sg.Pres build.ppt.FSg=OCL.3SgF all alone
 'I built it all by myself.'

Once again, then, we see significant variation in proclisis-enclisis options. These depend on left-peripheral features (C, *wh*, negation, focus) and features of the subject (definiteness, quantification); this is the European Portuguese type of system. The second set of options revolves around finiteness (with considerable variation in periphrastic tenses involving participles, etc.); this is the majority type, with much subvariation, including a major difference between Spanish/Italian and French regarding the basic pattern in infinitives.

Finally, clitics can be systematically enclitic, not just to the verb but also to postverbal negation and adverbs. We can characterize these four patterns as involving, respectively, (a) clitic-movement to the left-periphery, (b) clitic-movement to T, (c) clitic-movement to finite T only, and (d) clitic-movement to a lower head ($Asp_{\text{terminative}}$ according to Tortora 2014b: 100). In all these cases, clitic-movement interacts with verb-movement. In periphrastic tenses, as we saw, there is further variation, which may depend either on the position of the participle or on the height of clitic-movement (or possibly both), as well on the type of auxiliary and the type of clitic (this latter factor is also operative in Romanian). Finally, a further dimension of significant variation, not considered here, concerns the various kinds of ‘clitic-climbing’, i.e. placement of the clitic (either proclitically or enclitically) on a non-local verbal host (see Roberts 2016: 799–800 for an overview). Given their interaction with features of the left periphery and TP (finiteness, verb-movement), I take it that this variation is illustrative of further Romance-internal parametric variation.

1.2.4 Variation in past-participle agreement

The data presented here and in the next subsection are largely taken from Ledgeway (in press). In order not to unduly encumber the presentation, I refer just to Ledgeway’s paper; its references will lead the interested reader to his sources.

Ledgeway gives six possibilities as ‘a representative sample’ of variation in past-participle agreement across Romance. Restricting attention to active past participles, and assuming throughout that the agreement observed—which usually involves number and gender (not person)—is a manifestation of Agree in ϕ -features between v_{PTP} (participial ‘little’ v , located in the series of $v\text{P}$ -shells which, along with Voice, constitute the lowest clausal phase; see Chapters 5 and 6) and a nominal in its c -command domain, we thus have seven different feature specifications for v_{PTP} . A first option is represented by many Ibero-Romance varieties, which simply fail to show agreement (in the active; they do agree in gender and number with an internal nominal argument in the passive). This is illustrated by Spanish (Ledgeway’s (24), p. 11; here and elsewhere, both the example and the gloss have been insignificantly modified):

- (38) La manzana, la había comido. [Spanish]
 The.FSg apple it.FSg= have.1Sg.Pst eat.ppt.MSg
 ‘I had eaten the apple.’

Here the participle does not agree with the feminine singular direct object, but instead shows default masculine singular agreement.

A second option is shown by Ariellese, a Central-Southern dialect of Eastern Abruzzo in Italy, where the participle agrees in number with any plural internal or

(39) a. Seme magnite lu biscotte. [Eastern Abruzzese]
Be.1Pl.Pres eaten.MPl the.MSg biscuit.MSg
'We have eaten the biscuit.'

b. So magnite li biscutte.
Be.1Sg.Pres eat.ppt.MPl the.MPl biscuits.MPl
'I have eaten the biscuits.'

(Ledgeway's (24b), p. 11)

(40) Avètz presas de fotòs? [Occitan]
Have.2PL.Pres take.ppt.FPl of photos.FPl
'Did you take any photos?'

(41) la clé que j' ai prise [French]
 the.FSg key.FSg that I=have.1Sg.Pres take.ppt.FSg
 'the key which I took'

⁶ In Romance generally (with some complications in French and other Gallo-Romance varieties, as well as in Catalan; see n. 8), the surface exponence of past-participle agreement lies in the form of the final vowel. However, in Ariellese, as in many Central and Southern Italo-Romance varieties, the surface synchronic exponence of past-participle agreement is shown by the height of the stressed vowel of the participle, with final vowels reduced to schwa. This is the result of diachronic processes of metaphony, or vowel harmony, typically concerning the height or backness of the vowel. In these examples, the citation form of the participle is *magnate*, with the penultimate vowel the theme vowel of the first conjugation, as elsewhere in Romance. This vowel is raised as a synchronic reflex of MPI agreement, a diachronic reflex of height/backness harmony with the earlier MPI *-i* ending.

⁸ The exponence of past-participle agreement in French is often eliminated by phonological rules. In non-liaison contexts, plural *-s* undergoes the final-obstruent deletion pervasive in French (Dell 1985: 178ff.). Feminine *-e* (schwa) is rarely pronounced. However, where the participle underlyingly ends in an obstruent, as in the text example, the feminine schwa prevents final-obstruent deletion, giving rise to audible indirect exponence of agreement with the feminine form being /priz/ and the masculine /pri/. Similarly, in Catalan the absence of a vowel indicates MSG, while the presence of a vowel indicates feminine and with concomitant voicing effects on the final dental, e.g. *menjat* 'eaten' (MSG) vs *menjada* 'eaten' (FSG).

- In Sardinian (43a), there is the further restriction that participles can only agree with a 3rd-person direct-object clitic, and in Standard Barcelona Catalan (43b) the agreement is limited to 3rd-person feminine direct-object clitics (Ledgeway's (24f, 24g), p. 11):

- Ledgeway (in press: 12) arranges these options into a parameter hierarchy of the kind we will discuss extensively in §1.5 and subsequent chapters; see also Ledgeway & Roberts (2017: 602). Here, though, the purpose is simply to illustrate a further dimension of variation, this time presumably involving Number and Gender features of the head associated with the participle, Ledgeway's v_{Pt} .

The periphrastic perfect forms of Romance, which in many varieties, including Standard French and Regional Northern Standard Italian, is the unmarked past (see §5.3.1), are formed from the past participle of the lexical verb combined with either a reflex of Latin *esse* (BE henceforth) or a reflex of Latin *habere* (HAVE henceforth).⁹ The choice of auxiliary is, as Ledgeway (in press: 14) puts it, an ‘area of spectacular diachronic and synchronic microvariation’. There are three main dimensions of variation, which in many varieties interact in an intricate way: the tense or mood of the clause, the Person/Number specification of both the external and the (direct) internal argument, and the nature of the internal argument, particularly with unaccusative verbs where the internal argument is the sole argument of the verb. There are also degrees of free variation, the most extreme of which is ‘genuinely free variation between HAVE and BE in all contexts’ (Ledgeway, forthcoming: 14, n. 11), illustrated by the Southern Calabrian of Saline Ioniche.¹⁰

⁹ In Portuguese and elsewhere in Ibero-Romance a reflex of Latin *tenere* ('hold') is used for the perfect. I leave this aside here.

¹⁰ As in §1.2.4, I refer only to Ledgeway's paper, where his sources are fully documented.

English. Many Central and Southern Italo-Romance varieties generalize BE to all contexts; this is also what we find in South Slavic.

Mood determines auxiliary selection in Romanian, with HAVE appearing in Realis contexts and BE in Irrealis ones (Ledgeway's (30a, 30b), p. 15):

- (44) a. Am mâncat. [Romanian]
 Have.1Sg.Pres eat.ppt
 'I have eaten.'
- b. Vor fi mâncat.
 Will.3Pl be.Inf eat.ppt
 'They will have eaten.'

In the Italo-Romance Campanian variety of San Leucio del Sannio, tense determines auxiliary choice. Two perfect forms (one indicative, one subjunctive) show HAVE, while the pluperfect indicative has BE (Ledgeway's (31a, 31c), p. 16):

- (45) a. Èggio fatto tutto. [Campanian of S. Leucio del Sannio]
 Have.1Sg.Pres do.ppt all
 'I have done everything.'
- b. Èrem' auta dice quéllo.
 Be.1Pl.Pst have.ppt say.Inf that
 'We had had to say that.'

Person-driven auxiliary selection is found in many Central and Southern Italo-Romance varieties as well as in certain Northern Catalan dialects. A very common pattern is BE in the 1st and 2nd persons and HAVE in the 3rd persons, abbreviated BBHBBH. This is found in Ariellese, the variety of Eastern Abruzzese discussed in D'Alessandro & Roberts (2010):

- (46) So /si /a fatecate. [Eastern Abruzzese of Arielli]
 Be.1Sg/2Sg/have.3SgPres work.ppt
 'I/you/(s)he have/has worked.'

It is unclear whether there are any varieties which are purely sensitive to Number as opposed to Person (or a combination of the two); Ledgeway (in press:19, n. 16) mentions the Pugliese dialect of Ruvo di Puglia, which has H~B, BBHHH in the perfect (where '~' henceforth indicates free variation). Aside from in the 1Sg, this variety shows Number-driven auxiliary selection.

As is well known, argument structure determines auxiliary selection in Standard Italian, with HAVE being the auxiliary with transitive and unergative verbs and BE with unaccusatives:

- (47) a. Abbiamo fatto buon viaggio. [Italian]
 Have.1Pl.Pres make.ppt good journey
 'We had a good journey.'

- ¹¹ See n. 6 on the vowel-alternation in the participle.

Spanish, BE is found with unaccusatives in Realis tenses, but exclusively HAVE in Irrealis tenses (Ledgeway's (42b)):

- (49) Si el sieruo que es fuydo mora mucho en casa/si
 If the servant that be.3Sg.Pres flee.ppt stays much in house/if
 ladrones que furtan de dia & de noche ouissen entrado
 thieves that steal of day & of night have.3Pl.Subjunc enter.ppt
 [Old Spanish]
 'If the servant who has fled remains a long time at home/if thieves who steal
 by day and night had entered.'

A different kind of interaction is seen in the North-Western Catalan dialect of Pont de Suert, where BE appears with unaccusatives in the pluperfect only, and in the Campanian dialect of Procida (an island in the Bay of Naples) where HAVE is found with transitives and unergatives only in the present perfect, BE being generalized elsewhere.

Other varieties combine all three factors, showing the BBHBBH Person split with argument-structure sensitivity in the 3rd Persons of the present perfect only. The South-Eastern Marchigiano dialect of Ortezzano has this pattern, with BE generalized elsewhere, while the South-Eastern Abruzzese dialect of Tuffillo generalizes HAVE. The Pugliese dialect of Minervino Murge has free variation in all Persons except 3rd, where unaccusatives take BE and transitives/unergatives take HAVE. Another Pugliese dialect, Altamura, shows this pattern in the present perfect, but free variation throughout the pluperfect and a HAVE/BE contrast in the counterfactual conditioned by epistemic vs optative modality. Some varieties restrict argument-structure sensitivity to just the 3Sg of the present perfect (Zagarolo, Lazio; Secinaro, Abruzzo). The Molisan dialect of Guardaregia and the Abruzzese dialects of Castelvechio Subequo and Montenerodomo have Person splits in the Plural with no sensitivity to argument structure in the 3rd Person, and sensitivity to argument structure just in 3Sg. The Pugliese dialect of Poggio Imperiale has BBB in all verb classes in the singular and in the plural with unaccusatives, but H~B,B,H~B with unergatives and transitives. Further interactions of this kind, involving all the dimensions discussed so far, are found in the Molisan dialects of Guardaregia and Capracotta and the Basilicatese dialect of Miglionico; see Ledgeway (in press: 28, n. 31) for details and references.

A further range of dialects is sensitive to whether the single argument of an unaccusative is a reflexive or not. In the Rhaeto-Romance of the Lower Engadine, only reflexive unaccusatives occur with BE, HAVE being generalized elsewhere. In many Northern Italian varieties non-reflexive unaccusatives show BE but reflexive unaccusatives show free variation. In the South-Eastern Piedmontese dialect of Castellazzo Bormida, reflexive unaccusatives show a BBHBBH Person split; the Friulian variety of Grizzo is the same, but with a HBHBBH Person split with reflexive unaccusatives. Some varieties show this pattern, but restrict the Person split to the present perfect; this is found in the varieties of San Lorenzo del Vallo, Calabria, and Buonabitacolo, Campania (with HH(H~)BHHH).

Other varieties show an interaction of reflexive marking and tense. The Salentino dialects of Maglie and Giurdignano/Uggiano la Chiesa show BE with

non-reflexives and HAVE with reflexives in the present perfect, but generalized BE in the pluperfect and counterfactual. The Basilicate dialect of Accettura is similar, but allows free variation in the present perfect. Some varieties have distinct Person splits determined by reflexivity: the Abruzzese dialect of Popoli shows HAVE in the plural in all Persons and all verb classes, but in the singular non-reflexive unaccusatives show BBB, transitives and unergatives show BBH, and reflexive unaccusatives BB,H~B in the present perfect. In other periphrastic tenses, all verb classes have BBBHHB in the pluperfect and generalized HAVE in the counterfactual. These patterns are shown in (50) (Ledgeway's (58), p. 33; there is also gemination of the initial consonant of the participle following BE in the 1Sg and 2Sg forms, not shown here):

(50) Abruzzese of Popoli:

- a. (tʃ) avemmə/ (v) ave:tə/ (s) annə
 OCL.1Pl=have.1Pl.Pres/OCL.2Pl=have.2Pl.Pres/self=have.3Pl.Pres
 vəniutə /durmoitə/ʃʃakkwa:tə
 come.ppt /sleep.ppt/wash.ppt
 'We/you/they have come/slept/washed.'
- b. sɔ /ʃi /ε vəniutə.
 be.1Sg.Pres/2Sg.Pres/3Sg.Pres come.ppt
 'I/you(Sg)/(s)he have/has come.'
- c. sɔ /ʃi /a dduimoitə.
 be.1Sg.Pres/2Sg.Pres/3Sg.Pres sleep.ppt
 'I/you(Sg)/(s)he have/has slept.'
- d. mə sɔ /tə ʃi /s a
 OCL.1Sg=be.1Sg.Pres/OCL.2Sg=be.2Sg.Pres/self=have.3Sg.Pres
 ~ε ʃʃakkwa:tə
 ~be.3Sg.Pres wash.ppt
 'I/you(Sg)/(s)he has washed.'
- e. (m) e:və/ (t) i:və/ (s) ev:ə/ (tʃ) avava:mə/
 OCL.1Sg=be.1Sg.Pst/OCL.2Sg=be.2Sg.Pst/self=be.3Sg.Ps/OCL.1Pl=
 (v) avava:tə/(s) ivənə vəniutə/ durmoitə/ʃʃakkwa:tə have.1Pl.Pst/
 OCL.2Pl=have.2Pl/self=be.3Pl.Pt come.ppt/sleep.ppt/wash.ppt
- f. (m) avessə/ (t) evissə/ (s) avessə/
 OCL.1Sg=have.1Sg.Subjunc/OCL.2Sg=have.2Sg.Subjunc/self=have.3Sg.
 /(tʃ) avassa:mə/ (v) avassa:tə/ (s) avissənə
 Subjunc/OCL.1Pl=have.1Pl.Subjunc/OCL.2Pl=have.2Pl.Subjunc/self=have.
 vəniutə/durmoitə/ʃʃakkwa:tə
 3Pl.Subjunccome.ppt/sleep.ppt/wash.ppt

The Southern Campanian dialect of Padula also shows a three-way split in the singular of the present perfect: non-reflexive unaccusatives show H~B,B,H~B, reflexive unaccusatives HHB, transitives and unergatives HH,H~B.

Finally, there are Romance varieties in which BE is lexically restricted with unaccusatives. This is true of Modern Standard French, as is well known, where

clear unaccusatives such as *disparaître* ('disappear') take HAVE. Other Northern French varieties have still fewer unaccusatives taking BE; Ledgeway (in press: 34) cites the Picard of Nibas and Valenciennes as limiting BE to *mourir* ('die') and *aller* ('go'), and the Lorrain variety of Ranrupt as selecting BE only with the latter. There is also well-known sociolinguistic variation in this respect in Canadian French (Sankoff & Thibault 1977). Some varieties of Québécois show HAVE with unaccusatives denoting punctual events and BE with unaccusatives in a resultant-state interpretation (a similar phenomenon is observed in Romanian; see Ledgeway, in press: 34 for discussion and references).

Whatever the mechanisms responsible for auxiliary selection (for proposals, see Kayne 1993; Cocchi 1995; D'Alessandro & Roberts 2010), we are clearly witnessing intricate and impressive microvariation here. If F is the feature determining selection of HAVE (the analyses just cited all follow Freeze 1992 in taking HAVE to be the auxiliary associated with the more complex structure; it is also cross-linguistically more marked—see Haspelmath 1998; 2001), then F interacts with Person, Number, Tense, and Irrealis features, as well as reflexivity and unaccusativity properties of verbs. So, again, we see six interacting axes of variation. It should also be borne in mind that the above is almost certainly an oversimplification of the situation, as many varieties remain undescribed and many have been lost (see Loporcaro 2016: 816 on the number of possible systems, and Ledgeway & Roberts 2017: 602 for a possible parameter hierarchy).

1.2.6 A comparison with Japanese

The purpose of this section is not to provide a detailed comparison of the Romance languages and Japanese, but rather to show that, despite the intricate variation we have observed in the previous five subsections, from a wider typological perspective the Romance languages are all rather similar, and may in fact occupy a relatively small niche in the overall range of available variation. Japanese is chosen as it is a well-studied language which is typologically and genetically remote from Romance. The comparison brings out the macro-similarities among the Romance languages, as against the micro-differences we have been observing in the previous subsections. It also suggests that different properties vary to different degrees, as we shall see.

Before looking at what the Romance languages share, and undertaking the comparison with Japanese, let us briefly recap what we have observed in the foregoing. In §1.2.1, we observed that variation in SCL systems across Romance concerns, first, the presence of SCLs, whether the SCLs are heads, a cross-cutting set of options concerning the feature content and realization of at least four functional heads in the inflectional field and the left periphery, with the relevant features including ϕ -features, *wh*-features, and features sensitive to definiteness and/or quantification; in addition, the implicational scale in (17) represents five options, and there are three different left-peripheral licensing options and two different 'placeholder' options. In total, then, there appear to be ten binary options and a cross-cutting set of four feature combinations; treating these as independent binary options gives $2^4 = 16$ further options, i.e. 26 in all. In §1.2.2,

we observed three principal patterns of negation and two types of negative concord. In §1.2.3, we observed four types of clitic-movement (presumably implicating formal features of the left periphery, of T, of finite T only, or of a lower head, possibly $\text{Asp}_{\text{Terminative}}$). These options interact with the movement of finite and non-finite verbs in these fields (on which, see Schifano 2018 and §5.3.1), and there is a further option, barely discussed here, concerning clitic-climbing. Next, in §1.2.4 we saw seven options for past-participle agreement, presumably involving at least Number and Gender features of the participial head v_{Ptp} . Finally, in §1.2.5, we saw how auxiliary selection in periphrastic tenses is conditioned by Mood, Tense, Person, and Number features, with these options interacting in highly complex and intricate ways.

Let us, rather mechanically, attempt to get a sense of the range of possible variation all these interacting options can give us. There are 26 options involved in determining variation in SCLs, 3 types of clausal negation, 2 types of negative concord, 4 types of clitic movement, plus clitic-climbing, 7 options for past-participle agreement, and 4 interacting factors conditioning auxiliary selection. Taking the formal features involved in auxiliary selection to be binary and independent (and we saw good reasons for this assumption in the previous subsection), the number of possible variants will be 2^4 , i.e. 16. Further assuming that all the other axes of variation are independent, we have $26 \times 3 \times 2 \times 5 \times 7 \times 16$ possible grammatical systems, i.e. 87,560. Of course, any of our assumptions can be questioned, but, if anything, the dimensions of variation discussed and illustrated have been underestimated. For example, it is abundantly clear from the discussion in Ledgeway (in press) and his sources (notably Manzini & Savoia 2005) that there are very probably more than 16 different auxiliary-selection systems in Romance. The other figures are very likely to be underestimates. Twenty-six is almost certainly an underestimate of the number of different SCL systems; as mentioned above, Manzini & Savoia (2005: 1.371ff.) document 180 varieties. The question of how interconnected these axes of variation are has hardly been addressed, but it seems clear that most of these properties vary independently; robust implicational relations among them have not been proposed in the literature, and indeed Manzini & Savoia in particular argue that the range of microvariation in these and many other properties is vast, just in the Italo- and Rhaeto-Romance varieties they discuss. So 87,560 possible systems looks a very modest estimate of the range of overall possible variation for these properties across Romance. I am not claiming that there are 87,560 Romance languages (although this would not necessarily be wide of the mark; see the discussion in Kayne 2005a: 14), but simply that the variation we have described gives rise to that number of possible systems in principle.

But, of course, we know that the Romance languages, being a fairly closely related set of languages, are typologically quite homogeneous from a broader cross-linguistic perspective. That is, there are properties which we know to be cross-linguistically variant, but in relation to which the Romance languages are uniform or close to uniform. Let us consider a number of candidates in this connection.

First, the Modern Romance languages all show basic SVO word order and tend strongly to be harmonically head-initial (it is well-known that this property has

changed over the history of Latin/Romance; see Dankaert 2012; 2017; Ledgeway 2012). Of course, many of the world's languages are SOV, Japanese being a very well-known example. It is also well known that all the other logically possible basic orders of S, V, and O are found, although at radically different frequencies (see Dryer 2013c), SVO being the second most frequent order after SOV (488 SVO languages of 1,377 languages catalogued in *WALS*, i.e. 35.4% vs 565 SOV languages, i.e. 41.0%). Moreover, disharmonic word orders of various kinds are common, perhaps prevalent (see Biberauer 2008: 3; Biberauer and Sheehan 2013: 13–14; Cinque 2013a; Chapter 2).

Second, the Romance languages are non-ergative (although some of the patterns of auxiliary selection are reminiscent of certain types of 'split ergativity', especially, but not only, the Person splits; see Mahajan 1994 and Manzini & Savoia 2005: II.73off. for discussion). According to *WALS*, 32 of the 190 languages surveyed show ergative case-marking on full nominals, i.e. 16.4% (see also §6.3 and Sheehan 2017a; 2017b). In this respect, Japanese and the Romance languages do not differ, in that Japanese fairly clearly shows accusative alignment too.

Third, the Romance languages, with the possible exception of most creoles, are all inflectional languages showing a good degree of morphological fusion, especially in verbal morphology. Again, many languages seem to organize their morphological systems along different lines, with the classical typology (dating from the nineteenth century but influentially systematized by Sapir 1921) featuring agglutinating, analytic, and polysynthetic languages. Japanese is a well-known example of an agglutinating language.

Fourth, with the exception of some varieties of Brazilian Portuguese (see Munn & Schmitt 1999) and some creoles, no Romance language allows bare count nouns in argument positions. In other words, no Romance language allows the equivalent of *I saw cat/Ho visto gatto/J'ai vu chat* referring to a single instance of a cat (and, of course, English does not either). Japanese, on the other hand, allows this, as shown in (51):

- (51) Taroo-ga hon-o yonda. [Japanese]
 Taroo-NOM book-ACC read
 'Taroo read a/the book.' (Takahashi 2008: 408)

Fifth, as a probably related point (see Chierchia 1998; Longobardi 2008), all the Romance languages have definite and indefinite articles (see Vincent 1997 on the historical development of the articles in Romance; Vincent argues that this development is related to the development of object clitics, whose 3rd-person forms are homophonous with definite articles in many Romance varieties). Once more, this is far from universal: there are languages with just definite and no indefinite articles (e.g. Hebrew, the Celtic languages, Icelandic), languages with indefinite but no definite articles (e.g. Turkish), and languages wholly lacking in articles—Japanese and other East Asian languages, as well as many of the languages of Africa, are in this class. According to *WALS*, 387 of 620 languages surveyed (62.4%) have definite articles or affixes, while 195 (31.5%) have neither definite nor indefinite articles (Dryer 2013d). The possible importance of the

article system in determining other aspects of variation will emerge in Chapters 3 and 4 (see also Bošković 2008).

Sixth, again with the exception of innovative Brazilian Portuguese and a number of creoles, and with the status of French something of an open question (although see Roberts 2010b on this point), the Romance languages have moderately ‘rich’ agreement systems. It is well known that there are both ‘richer’ and ‘poorer’ languages in this respect: English is significantly poorer and East Asian languages lack verbal agreement marking altogether (although honorific marking should not be overlooked), while many head-marking non-configurational languages of the Americas and Australia have much ‘richer’ systems, in many cases marking all the arguments of the verb and with complex and significant interactions among the different kinds of agreement (see Jelinek 1984; Béjar and Rezac 2009; Bány 2018).

Seventh, as we have seen in §1.2.3, the majority of Romance languages, with the exception of a small number of Rhaeto-Romance varieties, some varieties of Brazilian Portuguese, and some creoles, have complement clitics. Once again, this is clearly a highly variable property. Certainly, Japanese has nothing of this kind.

Finally, the Romance languages as a whole fail to show a full morphological case system, although both Romanian (with systematic syncretic genitive-dative marking) and Old French and Old Occitan (with a nominative vs non-nominative distinction in one of the two declensions of masculine nouns) have what we can think of as ‘partial’ case systems. Again, of course, case inflection is cross-linguistically very common; WALS lists 100 out of 261 languages surveyed (38.3%) as having overt case-marking of some kind (Iggesen 2013). The Japanese particles *ga* (nominative), *o* (accusative), etc. are typically analysed as case-markers, and so Japanese is usually thought to have a case system (see Saito 2007; 2016; Chapter 2, n. 33, and §6.5 on this last point).

So we see that, comparing Japanese with Romance regarding these features which are shared by Romance, we find a range of significant differences: Japanese is SOV, it is agglutinating, it allows bare count nouns in argument positions, it lacks definite and indefinite articles, and it has overt case-marking. On the other hand, it is like Romance in having accusative alignment, as shown by the distribution of the case particles *ga*, *o*, *ni*, etc.

Of course, we can simply say that the Romance languages share some parameter values (word order, and whatever parameters determine the presence of articles, etc.), while others (those determining the phenomena seen in §§1.2.1–5) vary within the family. But if we compare Japanese with Romance regarding this latter set of phenomena, a different picture emerges. Let us now turn to this comparison.

Concerning subject clitics, enclisis of complement clitics, and past-participle agreement, the situation is straightforward: Japanese simply lacks all forms of pronominal clitics and agreement. Indeed, Japanese may lack pronouns altogether; the words often used to translate English words such as *I*, *you*, *he*, *they* (or, of course, their Romance equivalents), the *daimeishi*, are really a subset of nouns, differing from English/Romance-style pronouns in being able to be

modified and in not forming a closed class. Similarly, outside the honorification system, there is no discernible agreement marking of any kind. The Agree operation may be at work with certain mechanisms of subject- and object-honorification (see Boeckx and Niinuma 2004, but see Bobaljik and Yatsushiro 2006 for counterarguments), and Watanabe (2008: 535–6) proposes that Agree determines important aspects of DP-internal syntax in Japanese, but the features involved here are not the familiar ϕ -features which are clearly involved with clitics (both subjects and complements) and past-participle agreement in Romance.

Turning next to auxiliary selection, in §1.2.5 we summarized the impressive range of variation in Romance regarding this phenomenon. In order for auxiliary selection of any kind to exist, a language must have a periphrastic perfect. Modern Japanese does not have a periphrastic perfect; the morpheme *-ta* is usually said to be a past-tense marker but it does have some properties of a perfect marker (see Ogihara 1999). More generally, periphrastic perfects involving a grammaticalized transitive verb of possession are cross-linguistically quite rare, particularly outside Europe (partly because transitive verbs of possession, ‘have’-verbs, are rare outside Europe; see Drinka 2017). Haspelmath (1998; 2001) identifies this as a feature characterizing a putative European linguistic area. So nothing approximating to Romance auxiliary selection can be found in Japanese.

Japanese does, of course, have negation; indeed, clausal negation is a good candidate for a substantive universal in the sense of Chomsky (1965: 28), a notion we will propose a tentative definition of in §1.3. In *WALS*, Dryer (2013e: 1) notes, ‘There are no known instances of languages in which negation is realized by a change in word order or by intonation, and all languages have negative morphemes,’ while Talmy (2011: 755) considers the concepts ‘negative’ and ‘polarity’ to be what he calls ‘positive absolute universals’, i.e. (closed-class) elements found in every language. Furthermore, Bernini & Ramat (1996: 1) argue that ‘there is no known language which does not have some means or another of expressing negation’, and Willis et al. (2013: 1) say that negation ‘is one of the few truly universal grammatical categories’ (see also §7.3.1). The basic clausal negators in Japanese are the verbal suffixes *-nai/-masen* (the former informal, the latter polite), which attach to the verb stem, causing various morphophonological adjustments:

- (52) Takasi-wa ringo-o tabe-masen. [Japanese]
 Takasi-TOP apple-ACC eat-POLITE.NEG
 ‘Takasi does not eat apples.’ (Kishimoto 2008: 115)

The negative suffixes precede some verbal suffixes, e.g. the past-tense marker (showing conditioned allomorphy in this context):

- (53) John-wa choushoku-o tabe-nak-atta. [Japanese]
 John-TOP breakfast-ACC eat-NEG-PAST
 ‘John didn’t eat breakfast.’ (Watanabe 2004: 569)

But it follows other suffixes, e.g. *-rare-*:

- As the parallels between (ic–e) show, *-na* appears to be an adjectival suffix. Thanks to Mamoru Saito, p.c., for pointing this out to me. I will briefly discuss Japanese negation again in §§5.7 and 7.3.2.

acquisition: do we really want to say that there are 57 innately given options, all of which the Japanese child must try out against the PLD and set to negative? This is clearly a highly unattractive picture, in the sense that typological generalizations are being missed and (given the nature of parameters as simultaneously making claims about typology and acquisition) an implausible picture of acquisition is being offered. Furthermore, we are placing a heavy burden on UG.

The obvious solution to this problem is easy to see: the differences between Romance and Japanese that we have observed here show us that ϕ -features and periphrastic tenses are optional at the UG level, while negation is not. There may be a functional account for this, although I will not speculate on that here (see §§7.3, 7.4). In parametric terms, then, the simplest account of the data we have observed is to say that there is a single ' ϕ -parameter', which Japanese sets negatively (essentially this was proposed by Fukui 1986 and Kuroda 1988; Saito 2007; 2016 develops a version of it too, as we will see in §3.2.4). The variation in Romance concerning SCLs, pro- and enclisis, and past-participle agreement is a consequence of the fact that the ϕ -parameter is set positively across the whole of Romance; these phenomena represent sub-options within this general option, and the variation across Romance which we have observed represents further sub-options, some of them extremely intricate. Similarly, there may be a 'periphrastic-tenses parameter', although this seems less likely; it is more probable that periphrastic tenses result from a conspiracy involving various parameters (see the brief comparison of Romance and Bantu verb forms at the end of §5.3.1). The idea of sub-options within more general options implies that there are different kinds of parameters which fall into implicational relations. This is not a new idea; as we saw in §1.1.1, Chomsky (1981: 6–7) anticipated it. But in what follows we will develop this basic idea into a general theory of parameter hierarchies. Our Romance–Japanese comparison makes the basic *prima facie* empirical and theoretical case for doing this.

The conclusion that negation is not optional does not entail that it is not subject to parametric variation; in fact we observed variation in the position of negation across Romance in §1.2.2. But there is no 'Neg parameter'. Instead, there is a small number of basic options for the realization of clausal negation, as we will see in §7.3.2. This may be because negation itself is an innate category (see Horn 2001: 154ff. for extensive discussion); perhaps for this reason or perhaps for more general, 'functional' reasons, expressing negation may be a 'no-choice' parameter of the kind discussed in Biberauer, Roberts, & Sheehan (2014) and in §3.6.3 below (see also §7.3 on negation parameters, both no-choice and otherwise).

The idea of sub-options is further supported by another aspect of our Romance–Japanese comparison. We observed an enormous amount of intricate variation in Romance but, nonetheless, a number of properties are common to all (or almost all) Romance languages. Again, this suggests that not all parameters are equal, and that the hierarchies may show common 'zones' reflecting different types of option. Once again, this is important both for typology and for acquisition. For typology, it implies that we may need to think in terms of the 'granularity' of language types that we can set up. In acquisition terms, certain features must be more readily acquired than others. If syntactic change arises from