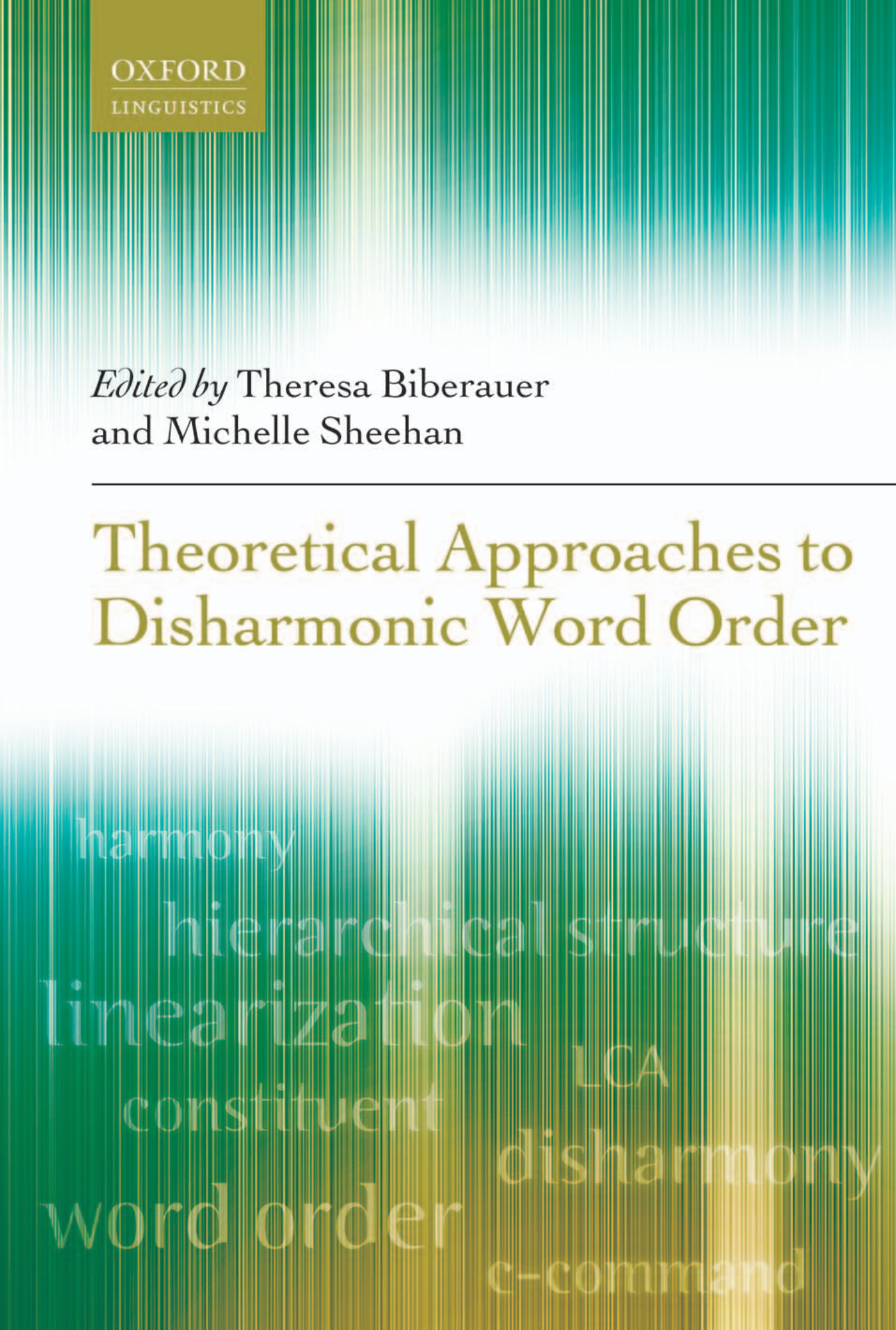

OXFORD
LINGUISTICS
Edited by Theresa Biberauer
and Michelle Sheehan
Theoretical Approaches to
Disharmonic Word Order
harmony
hierarchical structure
linearization
LCA
constituent
disharmony
word order
c-command

Theoretical Approaches to Disharmonic Word Order

Theoretical Approaches to Disharmonic Word Order

Edited by
THERESA BIBERAUER AND
MICHELLE SHEEHAN

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

First published in 2013

Impression: 1

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: 2013948770

ISBN 978–0–19–968435–9

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

Contents

Acknowledgments

This volume takes its name from the eponymous conference held at Newcastle University in 2009, organized by Anders Holmberg, Michelle Sheehan, and Laura Bailey. All but one of the chapters originated as a paper presented at this conference, which was held under the auspices of the Arts and Humanities Research Council (AHRC)-funded project, *Structure and Linearization in Disharmonic Word Orders* (AHRC Grant No. AH/E009239/1; Principal Investigator: Professor Anders Holmberg).

In addition to the authors, without whose co-operation this volume would, of course, never have been possible, we would like to thank the following for their reviewing assistance: Klaus Abels, Edith Aldridge, Linda Badan, Lawrence Cheung, Silvio Cruschina, Maximiliano Guimarães, Bill Haddican, S. J. Hannahs, Roland Hinterhölzl, Anders Holmberg, Jaklin Kornfilt, Iain Mobbs, Neil Myler, Johan Oosthuizen, Matt Pearson, Marc Richards, Norvin Richards, Ian Roberts, Kirill Shklovsky, Andrew Simpson, Luis Vicente, Irene Vogel, George Walkden, Wim van der Wurff, and Maria Luisa Zubizarreta. Additionally, we are very grateful indeed to the two anonymous OUP reviewers who reviewed the entire manuscript, not only for their very positive evaluations, but also for their extremely helpful suggestions. Thanks too to Jessica Brown for her help with various formatting tasks; to Leston Buell for his phenomenal copyediting; and to Anna Asbury and Jess Smith for their meticulous work at the proofs stage. Last, but certainly not least, our grateful thanks to OUP's Julia Steer for her immensely efficient, but always friendly and, where necessary, patient support; and to Kate Gilks for having overseen the final stages of the project.

Theresa Biberauer and Michelle Sheehan

Cambridge, July 2013

The Contributors

Michael Barrie is an Assistant Professor of linguistics at Sogang University in Korea. His main research interests are in generative syntactic theory, phrase structure, Antisymmetry, Distributed Morphology, and the Minimalist Program. He is interested in noun incorporation, Control, *wh*-movement, Northern Iroquoian, Algonquian, Romance, and Chinese. He is also interested in on-going language revitalization efforts of Iroquoian languages in the Six Nations community in Canada.

Theresa Biberauer is a Senior Research Associate at the University of Cambridge, where she is also a Fellow of Churchill College, and an Associate Professor Extraordinary at her South African *alma mater*, Stellenbosch University. Her research interests are principally in theoretical and comparative (synchronic and diachronic) morphosyntax, with Germanic generally and Afrikaans in particular being areas of specific interest. Her past work has focused on word-order variation, (null) subject phenomena, negation, information structure, and the larger question of the nature of parametric variation.

Brian Hok-Shing Chan is currently an Associate Professor at the University of Macau. He obtained his PhD in Linguistics at University College London, where he completed his PhD thesis entitled *Aspects of the Syntax, Production and Pragmatics of Code-Switching—With Special Reference to Cantonese and English*. A revised version of the thesis was published subsequently by Peter Lang in 2003. Since then he has continued his research on code-switching, while retaining an interest in the syntax of Cantonese. More recent publications can be found in *Lingua* (2008) and *The Cambridge Handbook of Linguistic Code-Switching* (2009) and *World Englishes* (2009).

Guglielmo Cinque is Professor of General Linguistics at the University of Venice, with interests in syntactic theory, typological linguistics and Romance linguistics. He has been involved in teaching and administration at the University of Venice and has taught at various universities abroad (Paris VIII, Vienna, Geneva, Harvard, UCLA, Ecole Normale, Paris, Brussels).

Since receiving her PhD in Linguistics in 2010 from the University of Padova, Federica Cognola has been working at the University of Trento as a Post-doctoral Research Fellow and a Contract Lecturer in German linguistics. Her research interests are syntax, monolingual and bilingual language acquisition, language variation and change, and the history of German and its dialects. She is the author of several articles on the syntax of the German dialect Mòcheno and of two monographs: *Acquisizione*

plurilingue e bilinguismo sbilanciato: uno studio sulla sintassi dei bambini mocheni in età prescolare (Padova: Unipress, 2011) and *Syntactic Variation and Verb Second: A German Dialect in Northern Italy* (Amsterdam: John Benjamins, 2013).

Redouane Djamouri is a Researcher at the Centre national de la recherche scientifique (CNRS) and affiliated to the Centre de recherches linguistiques sur l'Asie orientale (CRLAO) at the École des hautes études en sciences sociales (EHESS) and the Institut des langues et civilisations orientales (INALCO) in Paris. His main interest is the syntax of Early Archaic Chinese. He has published articles on various subjects concerning the language of the Shang and Zhou dynasties (13th–2nd c. BC): negation, deictics, focus clefts, prepositional phrases, and so on.

Arantzazu Elordieta is Associate Professor in the Department of Linguistics and Basque Studies at the University of the Basque Country. Her major areas of research interest include word order, null arguments, the syntactic and phonological realization of topic and focus, and syntactic microvariation within Basque. She is a research member of HITT (Basque Research Group of Theoretical Linguistics) and is also involved in the projects and activities organized by Basdisyn, a research group which brings together syntacticians, dialectologists, and computational linguists and studies syntactic variation phenomena across Basque dialects.

Joseph Emonds has published four books and numerous articles on syntactic and morphological analysis: *The Transformational Approach to English Syntax*, *Unified Theory of Syntactic Categories*, *Lexicon and Grammar: the English Syntacticon*, and *Discovering Syntax*. He is American, but moved to England in 1992. He has also taught in France, the Netherlands, Japan, Austria, and Spain, and now teaches and does research in the Czech Republic. His research focuses on how Universal Grammar interacts with formalized grammatical lexicons.

John A. Hawkins is Professor of Linguistics at the University of California Davis and Emeritus Professor of English and Applied Linguistics at the University of Cambridge. He has held previous positions at the University of Essex, the Max Planck Institute for Psycholinguistics in Nijmegen, and the University of Southern California, and visiting appointments at institutions including UC Berkeley, the Free University of Berlin, and the Max Planck Institute for Evolutionary Anthropology in Leipzig. He has broad interests in the language sciences and has publications in language typology, psycholinguistics, pragmatics, applied linguistics and second language learning, English and the Germanic language family, and historical linguistics.

Roland Hinterhölzl is Professor of German Linguistics at the University of Venice, Ca' Foscari. He has done research in comparative and diachronic German syntax. He has published an OUP monograph on scrambling, remnant movement, and restructuring in West Germanic (2006) and co-edited (with S. Petrova) a volume

on information structure and language change with Mouton de Gruyter (2009). He is currently working on issues of the interfaces between syntax, prosody, and information structure. His most recent publications include a comparative study of word-order regularities in German and English (*Syntax* 12, 2009), a paper on the development of V2 in Germanic (*Lingua* 120, 2011), and a lead paper (written with Ans van Kemenade) in the *Handbook of the History of English* (OUP 2012).

Richard S. Kayne has a PhD from MIT, a Doctorat ès Lettres from the University of Paris VIII, and honorary doctorates from the University of Leiden and the University of Bucharest. Currently Silver Professor at NYU, he is the author of *French Syntax: The Transformational Cycle*, *Connectedness and Binary Branching*, *The Antisymmetry of Syntax*, *Parameters and Universals*, *Movement and Silence*, and *Comparisons and Contrasts*.

Yasutomo Kuwana is a Senior English Teacher in charge of curriculum and instruction at Asahikawa Jitsugyo High School. He finished his master courses at Sapporo University and Ball State University in 2003 and 2005, and respectively obtained his Master's degrees in English Linguistics and Teaching English to Speakers of Other Languages (TESOL). He has contributed chapters in collected volumes published by Mouton de Gruyter. His research interests include phonology, typology, and second language acquisition.

Balkiz Öztürk is an Assistant Professor of Linguistics at Boğaziçi University, Department of Western Languages and Literatures. She received her Ph.D. degree from Harvard University in 2004. She has mainly worked on issues related to the syntax of Turkic Languages, such as relativization, incorporation, DP-structure, word order, EPP, case, and agreement systems. Her most recent research involves the interaction between event structure and argument structure in Laz, an endangered South-Caucasian Language spoken in Turkey.

Waltraud Paul is a researcher at the French Centre national de la recherche scientifique (CNRS) and affiliated to the Centre de recherches linguistiques sur l'Asie orientale (CRLAO) at the École des hautes études en sciences sociales (EHESS) in Paris. Her main interest is the syntax of modern Mandarin, with occasional excursions into the diachronic syntax of Chinese. She has published articles on a large variety of subjects: topic, focus, clefts, split CP, double object constructions, adjectival modification, serial verb constructions, and so on.

Michelle Sheehan is a Research Associate at the University of Cambridge specializing in comparative syntax with a particular interest in the Romance languages. She has worked on null arguments, Control, word-order variation, extraposition, clausal-nominal parallels, and case/alignment. She is co-author of the CUP volume *Parametric*

Variation: Null Subjects in Minimalist Theory and the forthcoming volumes *The Final-over-Final Constraint* (MIT Press) and *The Philosophy of Universal Grammar* (OUP).

Hisao Tokizaki is Professor of English Linguistics at Sapporo University. He finished the doctoral course at Hokkaido University in 1988 and received his PhD from Tsukuba University in 2007. A revised version of his dissertation was published in 2008 by Hituzi Syobo Publishers as *Syntactic Structure and Silence: A Minimalist Theory of Syntax–Phonology Interface*. He has published articles in journals including *English Linguistics* and *Phonological Studies* and contributed chapters in collected volumes published by Mouton de Gruyter and Equinox. His research interests include phonology, morphology, syntax, and typology.

Takashi Toyoshima received his doctoral degree from Cornell University in 2000. He is now an Associate Professor at Kyushu Institute of Technology, Japan. He is interested in the mathematical foundations of Minimalist syntax and has worked on the issues of computational complexity, derivational economy, phrase structure, and head movement, among other things. He has also co-authored a few papers in artificial intelligence and neural computation about induction of formal grammars with algorithms and in a self-organizing map.

Mark de Vos is a Senior Lecturer at Rhodes University in South Africa. He obtained an MPhil at Tromsø University in Norway before obtaining his PhD from the University of Leiden Center for Linguistics in the Netherlands. His research interests include the role of Dependency and Relational Theory in syntax, the interaction between syntax and PF output conditions, Agreement, pseudo-coordination, language variation, and, more recently, how linguistic-theoretical insights can be applied to indigenous-language literacies in South Africa.

John Whitman is Director of the Division of Crosslinguistic Studies at the National Institute for Japanese Language and Linguistics (NINJAL) in Tokyo and Professor of Linguistics at Cornell University. He works on historical/comparative linguistics, language typology, and syntactic theory, with a primary focus on the languages of East Asia.

List of Abbreviations

A	declension class in Urarina
ACC	accusative
ACT	active
AF	affective
AGR	agreement
AH	Accessibility Hierarchy
ALL	allative
AP	Adjective Phrase
ART	article
ASP	aspectual marker
AT	actor
AUX	Auxiliary (verb)
AuxP	Auxiliary Phrase
BBC	Basic Branching Constraint
BDT	Branching Direction Theory
BHR	Biberauer, Holmberg, and Roberts
BPS	Bare Phrase Structure
CCH	Cross-Category Harmony
CED	Condition on Extraction Domain
CL	clitic
CLF	classifier
CND	conditional/temporal
CNT	continuous aspect
COLL	collective
COMP/C	complementizer
CONJ	conjunction
CONT	continuative
COP	copula verb
CP	Complementizer Phrase
CRD	Constituent Recognition Domain
D	clension class in Urarina
DAT	dative

DEM	demonstrative
DET	determiner
DEVLD	devalued
DP	Determiner Phrase
DUR	durative aspect
E	declension class in Urarina
EIC	Early Immediate Constituents
EMPH	emphatic marker
EPP	Extended Projection Principle
ERG	ergative
EVID	evidential
EXP	experiential aspect
F	feminine
FD	functional dependency
FOC	focus
FOFC	Final-over-Final Constraint
FUT	future
FV	final vowel (Bantu)
GB	Government and Binding
GEN	genitive
H	head
HAB	habitual
HDT	Head Dependent Theory
HFF	Head-Final Filter
HRS	hearsay
IC	immediate constituent
iF	interpretable feature
IMPF	imperfective
IMPST	immediate past
IND	indicative
INDOBJ	indirect object
INGR	ingressive
IRR	irrealis
ITER	iterative
LCA	Linear Correspondence Axiom
LF	Logical Form

LOC	locative (verb)
M	masculine
MaOP	Maximize On-line Processing
MiD	Minimize Domains
N	neuter
NC	noun class
NEG	negation marker
NM	noun marker
NOM	nominative
NP	Noun Phrase
NSP	Natural Serialization Principle
NUM	Number
NUMCL	numeral classifier
NZLR	nominalizer
O	object
OSV	Object–Subject–Verb
OV	Object–Verb
OVS	Object–Verb–Subject
P	adposition (postposition/preposition)
p	light P
PART	particle
PASS	passive
PAST.HEARSAY	past evidential marking
PathP	Path Phrase
p^{DIR}	directional adposition
PFV	perfective aspect
PRF	perfect
PF	Phonological Form
PFV	perfective
PGCH	Performance–Grammar Correspondence Hypothesis
PL	plural
PLD	Primary Linguistic Data
P^{LOC}	locative adposition
POL	Polarity
POSS	possessive

POSSD	possessed
POST	postposition
PP	Adposition Phrase
PREP	preposition
PRES	present
PRET	preterite
PRO	pronoun
PROG	progressive
PTCP	participle
PST	past
PURP	purpose
PVC	postverbal constituent
Q	question particle/interrogative
QUAN	quantifier
REL	relativizer
Rel cl	relative clause
REM	remote past/future
S	subject
SFP	sentence-final particle
SG	singular
SOV	Subject–Object–Verb
SS	same subject marker
STR	strong
SUB	subordinating marker/subordinator
SUBJ	subject
SUP	superessive
SVO	Subject–Verb–Object
TOP	topic
TP	Tense Phrase (= IP– Inflectional Phrase)
uF	uninterpretable feature
UG	Universal Grammar
v	light verb
V	verb
V2	Verb Second
VO	Verb–Object

VOS	Verb–Object–Subject
VP	Verb Phrase
*v*P	light verb Phrase
VSO	Verb–Subject–Object
WALS	*World Atlas of Language Structures*
WI	Word Interpretation
1	first person
2	second person
3	third person

1

Theoretical Approaches to Disharmonic Word Order

THERESA BIBERAUER AND MICHELLE SHEEHAN

1.1 Introduction

Word order has not always been of great interest to grammarians. In ancient times, when the study of grammar meant the study of what we would today identify as the phonology, morphology, and syntax of a particular language,[1] word order was typically a minor syntactic concern, with largely morphologically based categorization considerations taking centre stage. As Henri Weil (1818–1909) notes, by the mid 19th century, grammarians had 'very carefully studied isolated words, as also their syntactical concatenation; but most of them [had] given no attention to the order in which words may follow each other' (1879 [1844]: 11). To the extent that they were concerned with word order, ancient grammarians were interested only in providing some rationale for the order of constituents. Thus Priscian (*floruit* 500 AD), drawing on the work of his contemporaries, proposed an abstract OV order for Latin based on the idea that 'the noun precedes the verb because the substance expressed by the noun precedes the accidents expressed by the verb' (Seuren 1998: 29, citing Luhtala 1994: 1467).

This pursuit continued into the 17th and 18th century, with grammarians being famously interested in word order as an indicator of the order of thought. Thus 18th-century linguists in the *grammaire générale* tradition compared the word orders of different languages in search of the *ordre naturel*. Some, including Nicolas Beauzée (1717–89), afforded SVO this status, making French 'analogical' in that its words tracked the order of thought, as opposed to Greek, Latin, and German, which were 'transpositive' as the correlation was indirect (Graffi 2001: 84). Others, including the

[1] As Seuren (1998: 29) notes: 'The Greeks and Romans were not directly concerned with universal properties of human language, their linguistic horizon being extremely restricted.'

philosopher Denis Diderot (1713–84), rejected the idea that the word order of a particular language could achieve such a status (cf. Graffi 2001: 17, citing Jellinek 1913–14 II: 425–64). The idea of a natural order nonetheless retained currency and is also observed in the work of Wilhelm von Humboldt (1767–1835) in the 19th century. Weil himself proposed a more nuanced status for word order, and he is perhaps most famous for noting that 'the syntactic march is not the march of ideas' (Weil 1879 [1844]: 21). This quote is misleading, however, as he nonetheless maintained that 'to treat of the order of words is then, in a measure, to treat of the order of ideas' (Weil 1879 [1844]: 11). Weil was principally interested in the reason why the modern European languages (French, German, English) had such little freedom in their word order compared with the Ancient languages (Latin and Greek). His claim in this connection is that, to the extent that word order is free, it reveals pragmatic meaning, an idea which later became central to the work of the Prague School and to which we return below.

Twentieth-century thinking about word order was initially strongly influenced by the views of Ferdinand de Saussure (1857–1913).[2] Given that the Father of Structuralism is often criticized for the relatively limited attention he paid to syntax as a component of language structure—certainly when compared to his consideration of phonology and morphology (cf. i.a. Joseph 2012: 540 ff. for discussion)—this may at first seem surprising. Saussure did, however, explicitly consider the nature of syntax generally and word order in particular in relation to his seminal distinction between *langue* and *parole* (see Belletti and Rizzi 2002: 1–4 for overview and Joseph 2012 for detailed discussion). For Saussure, the regularities of phrase construction clearly fell into the domain of *langue*; the freedom with which speakers are able to combine elements taking into account the discourse situation in which they find themselves, on the other hand, was for him *le propre de la parole* (Saussure 1916: 172, cited in Belletti and Rizzi 2002: 3), i.e. the domain of *parole*. Since word-order choices within a given language are to such a large extent conditioned by communicative considerations which individual speakers must weigh up, it is not difficult to see why following generations of linguists would have interpreted the minimal discussion in Saussure's *Cours*[3] as (further) justification for investigating word order primarily from a functional perspective. As such, Saussure's syntactic legacy contrasts sharply, and in a way that is not always recognized, with the influence his work had on phonology and morphology.

A major factor in Saussure's sparse discussion of word order having the influence it did within the domain of early 20th-century word-order research is undoubtedly

[2] We thank an anonymous OUP reviewer for drawing our attention to Saussure's influence on the direction that word-order studies took in the first half of the 20th century.

[3] As the OUP reviewer mentioned in the previous footnote points out, Saussure himself did no syntactic analysis.

also the work of Vilém Mathesius (1882–1945). Independently of Saussure, the founder of the Prague Linguistic Circle had also in his (1911) paper, *O potenciálnosti jevů jazykových* ('On the potentiality of language phenomena'), pointed to the distinction between the two forms of language that Saussure made famous as *langue* (cf. the *potenciálnosti* of Mathesius's 1911 title) and *parole*. Further, influenced by Masaryk (1885), he had also discussed the importance of a distinction betweeen 'static' and 'dynamic' aspects of language, one which he later related to Saussure's synchrony/diachrony dichotomy (cf. Mathesius 1927/1983). Unlike Saussure and the first generations of structuralists more generally, Mathesius, however, had firmly comparative linguistic interests, and his own work on Czech, German, and English reinforced in him the view that the appropriate *tertium comparationis* in comparative linguistics should be language function. More specifically, Mathesius's view was that languages differ in the structural means (word order, intonation, use of specific constructions, etc.) via which they permit speakers to communicate successfully, with the extent to which they draw on these structural possibilities defining their 'linguistic characterology' (cf. Mathesius 1928). He introduced to functionalist linguistics and the study of word order more generally the notion of 'functional sentence perspective',[4] in terms of which utterances can be divided into what is today referred to as a *theme* or *topic* (roughly, what the utterance is about) and *rheme* or *focus* (approximately, what is said about the theme/topic). Crucially, he also highlighted the way in which these notions correlate with what is today referred to as *information structure*: typically, the theme/topic maps onto discourse-old/salient information, while the *rheme/focus* corresponds to discourse-new/non-salient information. Importantly, this comparatively inspired work naturally connected with research, also being completed during the first decades of the 20th century, that was clearly inspired by Weil's earlier ideas. Based on his meticulous study of the history of German, Otto Behaghel, in his monumental four-volume *Deutsche Syntax: Eine geschichtliche Darstellung* ('German syntax: a historical account', published between 1923 and 1932), postulates a number of information-structure-sensitive word-order principles or Laws, which he assumed to be cross-linguistically valid. These include the following:

(1) a. **Behaghel's Second Law**: That which is less important (or already known to the listener) is placed before that which is more important (or unknown).

b. **Behaghel's Law of Increasing Terms** (*Gesetz der wachsenden Glieder*):[5] Given two phrases, the shorter precedes the longer where possible.

[4] More accurately, Mathesius himself employed the Czech term *aktuální členění větné* (literally 'actual division of sentence'), and it was Firbas (1957) who, building on Mathesius's own German translation (*funktionale Satzperspektive*), proposed the term *functional sentence perspective*.

[5] Following Cooper and Ross (1975), this Law is also often referred to as *Pānini's Law*.

During the early decades of the 20th century, then, various strands of European linguistic research were converging on the centrality of what would today be labelled 'functional' considerations in the study of word order, while simultaneously emphasizing the value of comparative work.

Strikingly, contemporaneous American structuralism was very different in its orientation, focusing on the purely synchronic structural description of individual languages and, in line with Bloomfield's (1934: 36) sentiments, avoiding the 'larger synthesis' or 'General Grammar, which will register the similarities between languages' until more was known about languages, non-Indo-European ones in particular (cf. also Bloomfield 1933: 46). As DeLancey (2001) notes, American structuralism's strong description-first/explanation-later orientation and the wider intellectual influence of behaviourism in psychology and logical positivism in philosophy created an intellectual climate within which comparative research received little attention. Thus early 20th-century European works such as those mentioned above, Wilhelm Schmidt's ground-breaking (1926) study of cross-linguistic variation in word-order patterns and their significance for language classification,[6] and even Sapir's (1929) pioneering classification of the indigenous languages of the Americas aroused little immediate interest in mainstream linguistics. It would take the work of Joseph Greenberg, student of both Sapir's teacher, Franz Boas, and, later, the Prague Circle's Roman Jakobson to unite the European and American research traditions and truly ignite 20th-century research in cross-linguistic word-order variation. In what follows, we review Greenberg's work and the influential typological tradition to which it gave rise before considering the status of word order in the generative tradition and future prospects in this domain more generally.

1.2 Harmony and disharmony from Greenberg to the present

1.2.1 *Greenberg's correlation pairs and the notion of harmony*

Greenberg (1963: 60) notes that 'linguists are in general familiar with the notion that certain languages tend consistently to put modifying or limiting elements before modified or limited, while others just as consistently do the opposite'. This is the basis of the notion of 'harmony' which lies at the heart of this volume. Thus, in English, modifiers tend to follow modified elements like verbs and adpositions, whereas the opposite is true in Hindi:

[6] In relation to Schmidt's work, it is worth noting that, despite its systematic treatment of word-order phenomena, its author's objective was not primarily linguistic; instead, it was intended as a vehicle for the interpretation of cultural history. Greenberg (1963: 105, note 4), who acknowledges the value of Schmidt's contribution, says of this 'applied' component of the work, 'His results there verge on the fantastic.'

(2) John [is [at [school]]] V P O

(3) Raam [[[skuul] par] hai] O P V
Raam school at is
'Ram is at school'

The basic word-order phenomena discussed by Greenberg are the following (based on Greenberg 1963: Appendices I–II):

(4) i. verb-initial/medial/final
ii. adposition–noun order
iii. noun–adjective order
iv. noun–genitive order
v. noun–demonstrative order
vi. noun–numeral order
vii. pronominal–verb order

For Greenberg, though, the term 'harmony' actually has a more technical definition, relying crucially on his notion of 'dominance'. A particular order is dominant over another order, where it is less constrained in the following terms:

> A dominant order may always occur, but its opposite, the recessive, occurs only when a harmonic construction is likewise present. (Greenberg 1963: 62)

This is effectively illustrated by a tetrachoric table, as shown in Figure 1. 1:

	V-DP	DP-V
pro-V	Y	Y
V-pro	Y	N

FIGURE 1.1 Positioning of lexical nominal and pronominal complements in relation to the verb

Figure 1. 1 indicates that three of four potential word-order combinations involving verbs and lexical/pronominal complements are attested. Crucially, the fourth, whereby pronominals follow, but full lexical complements precede the verb, is unattested. This means that V–DP is dominant over DP–V, as the latter order is more constrained than the former, occurring only where the order of pronominal and verb is harmonic with the order of full nominal (DP) and the verb.[7] Greenberg claims that harmony between two correlation pairs arises wherever we see this

[7] Note that this distributional fact might be viewed as support for Kayne's influential Antisymmetry hypothesis, discussed below and in several of the papers in the volume. Kayne (this volume) himself cites this universal in support of Antisymmetry.

particular pattern in a tetrachoric table. Note that for Greenberg, then, harmony is a notion which is defined across languages based on patterns of attested cross-linguistic variation (in this case a core sample of 30 and an expanded sample of 142 languages, but ideally the sum of all attested languages).[8] Interestingly, although the focus of his (1963) work was not primarily on explanation, his concluding speculations on this point did clearly indicate that harmony seemed to him 'very obviously connected with the psychological concept of generalization' (Greenberg 1963: 62; see also Hawkins 1980, 1982, and 1983 for further discussion and elaboration).

Many of the papers in this book adopt a slightly different notion of harmony, defined in relation to a specific language, but drawing on the cross-linguistic patterns first identified by Greenberg. In these terms, two phrases are harmonic if and only if their respective 'modifying or limiting elements' pattern together in *uniformly* occurring at the right or left edge of the phrase concerned. Thus VP and PP can be said to be 'harmonic' in this sense in the English and Hindi examples above: in both cases, the 'modified' V/P systematically occurs to the left/right of its 'modifier'. Likewise, in a specific language with pro–V and V–DP orders, the order of pronominals and verbs can be said to be disharmonic with the order of full DPs and the verb, as pronominals and full DPs (the 'modifiers' in this case) do not uniformly align left/right. This perspective on harmony naturally lends itself to interpretation in terms of a Head Parameter, as we shall see in section 1.2.2 below; more generally, it is also obviously compatible with Greenberg's (1963: 62) psychological generalization proposal, with Hawkins' much-discussed (1983) Principle of Crosscategorial Harmony, and with Roberts' (2007b) Input Generalization (see again section 1.2.2, and also section 1.3.1 below). Worth noting here, however, is that proposals to relate typological harmony and typological patterns more generally to Universal Grammar (UG) have been strongly challenged by both generativists and non-generativists (see i.a. Newmeyer 2005a, Haspelmath 2008b, Whitman 2008, and Boeckx 2010 for discussion). Non-generativists typically point to the paucity of genuinely exceptionless cross-linguistic patterns, calling into the question the role of uniformity-imposing UG. The typical generativist objection, in turn, is formulated by Newmeyer (2005a: 105) as follows:

[8] Greenberg's (1963) paper also arguably contains the roots of Cartography (cf. Cinque 1999, 2005a amongst many others). Consider the following:

> Another type of relation than those that have just been considered is illustrated by Universals 20 and 29. These may be called proximity hierarchies. What we have is a rule that certain elements must be closer to some central element than some other satellite. The central element may be the root morpheme or base of a word or the head-word of an endocentric construction. (Greenberg 1963: 104)

He goes on to add that '[t]hese hierarchies are presumably related to degrees of logical and psychological remoteness from the center, but no analysis is attempted here.'

Our minds/brains, after all, have no clue as to the typological status of any aspect of any element of our mental grammars. The relationship between typological generalizations and I-language is therefore necessarily quite indirect.[9]

And to this, we can also add Odden's (1988: 461) caveat that '[i]t is misguided to attribute every accidentally true statement about human language to UG, for doing so trivializes the theory of UG itself'.

We return to the issue of the relationship between (dis)harmonic word orders and mental grammars below (see the discussion in sections 1.3 and 1.4 in particular). Our immediate concern, however, is a more detailed consideration of the notion of 'harmony' and the theoretical ideas it has given rise to.

1.2.2 *Harmony and the Head Parameter*

Building on Greenberg's landmark paper, Vennemann (1972, 1974a,b) proposed an assimilation of the various word-order correlation pairs to a single operator/operand template. This in turn allowed him to posit a *Natural Serialization Principle* (NSP) whereby an unordered set {operator {operand}} is universally 'serialized' or linearized as either operator–operand or operand–operator in a given language. In later work (Vennemann 1976; Vennemann and Harlow 1977), Vennemann explicitly refers to operands as *Heads* and operators as *Specifiers*. As Dryer (1992a: 88) notes, however, it is important not to equate this use of *Specifier* with the generative notion 'specifier', as it is clear that Vennemann's intention here was to refer to dependent elements. The pairs in (5) illustrate the types of elements Vennemann was concerned with:

(5)

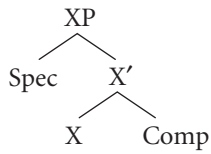
Operand/Head	**Operator/Dependent ('Specifier')**
Verb	Object
Verb	Adpositional Phrase
Verb	Manner Adverb
Noun	Relative Clause
Noun	Genitive
Noun	Adjective

In a sense, the NSP can be considered a precursor of the Head Parameter, the idea that in a given language L, a head universally precedes/follows its complement. Consider this in relation to Chomsky (1970) and Jackendoff's (1977) X-bar theory:

[9] Lightfoot's oft-cited (1979) objection to the Sapirian notion of 'drift' makes the same point in relation to diachronic typology:

> Languages are learned and grammars constructed by the individuals of each generation. They do not have racial memories such that they know in some sense that their language has gradually been developing from, say, an SOV and towards an SVO type, and that it must continue along that path. After all, if there were a prescribed hierarchy of changes to be performed, how could a child, confronted with a language exactly half-way along this hierarchy, know whether the language was changing from type *x* to type *y*, or vice versa? (Lightfoot 1979: 391)

(6)



Here, Vennemann's operand could be seen as equivalent to the Head X, while a relevant subset of his operators are equivalent to Comp.

Hawkins (1980) offers an eloquent critique of the NSP, based partly on the observation that only 47.89% of the languages in Greenberg's sample actually have consistent operator–operand/operand–operator order across all the categories he considered. These are given below:

(7) i. verb-initial/-medial/-final (collapsed to verb–object order by Vennemann—see (5) above)
ii. adposition–noun order
iii. noun–adjective order
iv. noun–genitive order[10]

The lack of ordering consistency across pairs of elements means that very few languages are consistently 'harmonic' in the second, consistent left/right alignment sense discussed in section 1.2.1, i.e. for a given system, the modified X systematically occurs in a fixed position—left/right—in relation to its modifier. As Hawkins notes, then, what might be termed *disharmonic languages* actually outnumber their consistently harmonic counterparts, and the NSP's inability to account for the former is thus a major failing.[11] Vennemann (1975) had in fact proposed that languages which fail to conform to either word-order type are in a state of flux, being in the process of undergoing a diachronic change, but, as Hawkins notes, this is a problematic claim. Of the 24 logically possible word-order combinations, eight remain unattested (cf. Hawkins 1980: 198).[12] There is thus a second sense in which the NSP is problematic, as it fails to provide an account of these eight unattested types. All orders diverging from the consistently harmonic orders (again, in the second, non-Greenbergian sense) have the dubious status of intermediary stages of diachronic change for Vennemann, and, unless independent considerations are identified as to why the unattested systems fail to surface as intermediary systems, they cannot be ruled out by his approach. It is also worth noting that many attested disharmonic systems do not obviously appear to be in the process of change in the direction of harmony, i.e. a

[10] Given Greenberg's original three-way distinction for (i), this gives 3×2^3 (24) potential combinations.

[11] As Hawkins puts it (1980: 198), a scientific theory which accounts for less that 50% of the data is not a good theory.

[12] Vennemann proposes the merging of SVO and VSO languages to give the single type VO, but Hawkins takes issue with this move, mainly because, as Greenberg (1963) showed, VSO languages display much stronger correlations than SVO languages.

stable system. English, for example, has retained its disharmonic Saxon genitive over many centuries, with this structure in modern English being significantly more productive than the postnominal PP option (consider: *my friend's house* and *the house of my friend*). Sapirian 'drift'-oriented interpretations of Vennemann's NSP therefore do not appear to hold up (cf. Sapir 1921 on 'drift' and Roberts 2007b: 340–57 for discussion of potential theoretical interpretations of this notion). Finally, Hawkins points out that the NSP posits a series of bilateral relations between correlation pairs which in turn prevents one from capturing Greenberg's notions of dominance and harmony in their original sense: unilateral implicational universals lie at the heart of the Greenbergian notion of dominance and, in eliminating them, Vennemann again detracts from the viability of the NSP.[13]

There is thus a sense in which the Head Parameter (in the form of the NSP) was proposed early in the typological literature and quickly rejected. In a slightly later paper, though, Hawkins (1982) returns to the issue and discusses the notion of harmony more explicitly in relation to X-bar theory and potential 'spec' and 'head parameters'. He observes that X-bar theory, with its three-way head/specifier/complement distinction, might provide a finer-grained distinction than the two-way modifier/modified or operator/operand distinctions used by Greenberg and Vennemann, respectively. Once again, though, he notes that there is no evidence that specifiers, heads, and complements are systematically ordered (albeit in potentially different ways) across a given language. Even once we allow for the specifier/complement distinction, he argues, 'languages will vary according to the degree of cross-categorial generalization which their grammars incorporate' (Hawkins 1982: 9). Translating into a generative perspective, then, Hawkins (1982) concludes that word order cannot be regulated by a single parameter; it might, however, be regulated by a series of spec and head parameters which are psychologically related. In fact, Hawkins argues explicitly that '[g]rammars with more cross-categorial generalizations will be simpler than, and hence preferred over, those with fewer'. This is essentially a formalization of Greenberg's intuition that harmony is connected to generalization. In the modern generative context, specifically, Chomsky's 'three-factors' framework (cf. Chomsky 2005), this can very naturally be understood as the consequence of an intuitively plausible 'third factor'[14]—something like Roberts'

[13] Hawkins goes on to propose his own theory of the Greenbergian word-order correlations based on the combination of four absolute implicational universals and his relative principle of 'Cross-Category Harmony' (CCH), which aims to predict the relative frequencies between the 18 permitted word orders. We return to this principle in section 1.3.1.

[14] In the context of Chomsky's 'three-factors' approach, the factors assumed to play a role in determining the form of adult grammars are specified as UG (Factor 1), the PLD (Factor 2), and, additionally, rather vaguely specified 'third factors' or non-language-specific considerations, which include principles of efficient computation and principles of data analysis employed in acquisition (see Mobbs 2008, in progress for further discussion).

(2007b) Input Generalization, a strategy acquirers are assumed to employ in analysing the Primary Linguistic Data (PLD) they are exposed to. Importantly, this 'third-factor' take on the source of harmonic patterns removes the need to appeal (stipulatively) to (first-factor) UG, while also potentially addressing generative concerns (such as those of Newmeyer (2005a) and Lightfoot (1979), highlighted in the previous section) about the feasibility of understanding typological patterns (here: consistent left/right alignment harmony). Particularly worth noting in the present context, though, is the fact, typically overlooked by generativists, that, as early as the early 1980s, Hawkins rejected the idea that a single word-order parameter (e.g. the Head Parameter) could account for attested word-order variation and also that this variation could be understood as a direct reflex of UG alone.

Dryer (199-2a) tests Greenberg's word-order correlations on a much larger, more balanced 625-language sample. One of the theoretical objectives of his research is to test the feasibility of what he calls *Head Dependent Theory* (*HDT*), stated in (9):

(9) **The Head Dependent Theory** (**HDT**)
Verb patterners are heads and object patterners are dependents, i.e. a pair of elements X and Y will employ the order XY significantly more often among VO languages than among OV languages if and only if X is a head and Y is a dependent (Dryer 1992a: 87).

In the context of a Head Parameter-based approach, the HDT, then, predicts that heads will be verb patterners, while complements are object patterners. Thirteen of the verb–object patterners identified by Dryer arguably involve head–complement relations. Consider (10):

(10) **Head–complement correlation pairs** (taken from Dryer 1992a: Table 39, 108)
(i) verb–object
(ii) adposition–DP
(iii) copula verb–predicate
(iv) want–VP
(v) auxiliary verb–VP
(vi) negative auxiliary–VP
(vii) complementizer–S(entence)
(viii) article–noun
(ix) plural word–noun
(x) noun–genitive
(xi) adjective–standard of comparison
(xii) verb–PP
(xiii) adverbial subordinator–S(entence)

The remaining four, however, are less obviously head–complement relations:

(11) **Other correlation pairs** (taken from Dryer 1992a: Table 39, 108)
  (i) question particle–S(entence)
  (ii) noun–relative clause
  (iii) verb–manner adverb
  (iv) verb–subject

Question particles are often taken to be heads selecting clausal complements (cf. Cable 2010 for recent discussion), but some problems for this are raised by Biberauer, Holmberg, and Roberts (in press) and Sheehan (this volume). In traditional analyses, relative clauses are taken to be adjuncts, but on Kayne's (1994) raising account, they are complements of the determiner, and so might arguably pattern with complements for this reason, as articles are verb patterners. Manner adverbs are generally assumed to involve either adjunction (cf. Ernst 2002) or, from a Cinque (1999) perspective, a specifier relation, and so are more problematic. Likewise, subjects are accepted to occupy a specifier position in X-bar theory, and so also present a potential problem. Evidently, then, Dryer's findings empirically reinforce Hawkins' observation that it cannot be a single Head Parameter which determines the observed harmonic patterns.

This point is also suggested by the potential correlation pairs (posited by Greenberg or others) which fail to pattern with verb–object order in Dryer's larger sample. Consider (12) in this connection:

(12) **Non-correlation pairs** (Dryer 1992a: 108, Table 40)
  (i) adjective–noun
  (ii) demonstrative–noun
  (iii) adverbial intensifier–adjective
  (iv) negative particle–verb
  (v) aspect/tense particle–verb

Worth noting about these pairs is that none of the non-correlation pairs involve clear head–complement relations: (i) is generally taken to involve adjunction or a spec–head relation in the extended nominal projection (Svenonius 1993), but not usually complementation (although see Abney 1987); (ii) is commonly thought to involve a spec–head relation (see i.a. Cinque 1995; Bernstein 1997 and 2008, and Giusti 2002); (iii) is taken to involve either adjunction or a spec–head relation (Bresnan 1973; Jackendoff 1977; though see Corver 1997 for a head analysis of a subset of degree words). The status of (iv) and (v) is unclear, as is made manifest by several of the papers in this volume (cf. Chan and Sheehan, and see also Biberauer and Sheehan 2011 on the problems posed by particles in the generative context). It is therefore not immediately clear that the non-correlating behaviour of the elements in (12) constitutes a challenge to the HDT. Dryer (1992a: 108–18), however, rejects

HDT-type, and thereby also Head Parameter-based, approaches, proposing the Branching Direction Theory instead:

(13) **The Branching Direction Theory (BDT)** (Dryer 1992a: 109)
Verb patterners are non-phrasal (non-branching, lexical) categories and object patterners are phrasal (branching) categories, i.e. a pair of elements X and Y will employ the order XY significantly more often among VO languages than among OV languages if and only if X is a non-phrasal category and Y is a phrasal category.

In terms of BDT, the non-correlation pairs in (12) can be understood as cases involving pairs of items which do not consistently exhibit an identifiable branching direction because (a) each of the two items in question can readily be represented by single words, e.g. *blue skies* (i), *this weekend* (ii), *very tall* (iii), *not leave* (iv), and *repeatedly coughing* (v), and (b) the modifying element in each case is not fully recursive in the sense that it can embed other XPs (e.g. PPs, NPs, or clauses). The difficulty with the non-correlation pairs in (12), then, is that the modifying element does not exhibit the expected phrasal properties. What is predicted, however, is that, where an optionally branching modifier is fully recursively phrasal in the above sense, it will exhibit the behaviour of an object patterner. This prediction is borne out in cases such as those illustrated in (14–15) below:[15]

(14) a. **blue** skies
b. *skies **blue**
c. skies [$_{AP}$ **blue** [$_{CP}$ **as the most brilliant sapphire**]]
d. ***blue as the most brilliant sapphire** sky

(15) a. J' admire **souvent** le courage de mon père
I admire often the courage of my father
'I often admire the courage of my father.'
b. *J' admire le courage de mon père **souvent**
c. J' admire le courage de mon père **quand je regarde à la télé des films sur la Deuxième Guerre mondiale**
I admire the courage of my father when I look at the TV of films over the second war world
'I admire the courage of my father when I watch TV programmes about the Second World War.'
d. *J'admire **quand je regarde à la télé des films sur la Deuxième Guerre mondiale** le courage de mon père

[15] Thanks to an anonymous reviewer for the examples in (15).

A clear difference between BDT and HDT, then, is that the former, but not the latter—and, by extension, therefore, also not Head Parameter-based approaches more generally—can account for the fact that harmony affects not only heads and their complements (defined in terms of subcategorization), but also potentially branching adverbial modifiers: in BDT terms, all 'fully recursive phrasal dependents' are expected to exhibit harmonic behaviour in relation to verb patterners (Dryer 1992a: 116). An anonymous reviewer suggests that BDT's ability to account for the placement difference between fully recursive and non-fully recursive modifiers renders it superior to accounts referring only to grammar-internal considerations. While we do not dispute this point (see below), we do, however, wish to note that there are data which pose a challenge to Dryer's BDT proposal (rejected in Dryer 2009).[16] These include phenomena such as that illustrated in (16):

(16) a. den över sin dotter stolt- a mamma-n [Swedish]
the of her daughter proud-DEF mother-DEF
'the mother who is proud of her daughter' (Cabredo Hofherr 2010: 15)

b. ett sedan i går välkänt faktum
a since yesterday well.known fact
'a fact well-known since yesterday' (Delsing 1992: 25)

Here a fully recursive AP modifier precedes the modifiee, despite the fact that object patterners in Swedish should follow their modifiees. Similar patterns are observed in a range of languages with otherwise head-initial nominals, i.e. nominals in which (fully recursive) dependents should, in BDT terms, follow their modifiee (see Sheehan 2012 for overview discussion and references). Dryer's processing-oriented account, then, also does not straightforwardly account for the observed harmonies and disharmonies. Importantly, Dryer's (1992a) and subsequent research finds that the preference for harmony (i.e. consistent patterning across verb and object patterners) is statistical rather than absolute, with very few languages emerging as fully harmonic in Dryer's terms (cf. Dryer 1992a: 109, note 17). To the extent that parameters can play a role in the understanding of harmony and disharmony, then, what seems to be required is a series of semi-independent parameters and, additionally, some overarching and quite possibly 'externally' (i.e. non-UG-imposed) preference for harmony (of the second, consistent left/right alignment type discussed above). In what follows, we will mostly restrict our focus to the disharmonic word

[16] Dryer's (2009) rejection of BDT is motivated by his rejection of the hierarchical constituents it assumes; in place of the BDT, he argues that the observed word-order patterns—some of which he shows, on the basis of a further enlarged sample, to hold even more strongly than was possible in the (1992a) paper—can be ascribed to more general processing considerations, the nature of which is, however, largely left to future research.

orders that are this volume's main concern and that Dryer's research in particular has shown to be cross-linguistically very common.

1.2.3 *Disharmony*

Today, there is recognition in both the typological and the generative literature that very many and possibly even the majority of languages fail to be fully harmonic in the sense that all head–complement pairs pattern alike (this point is once again picked up on by Cinque, this volume).[17] As an across-the-board Head Parameter, set once for all categories, clearly cannot account for the observed variation, generative grammarians have proposed interacting parameters designed to account for disharmonic word orders. Li (1990: 41), for example, proposes the following constraint to account for word-order patterns in Mandarin:

(17) **The Chinese Word-Order Constraint**

a. Chinese is head-final except under the requirements of Case assignment.

b. Case is assigned from left to right in Chinese.

c. A Case assigner assigns at most one Case.

This constraint has the advantage of accounting for the unusual word-order properties of Mandarin (e.g. the initial position of Case-assigning verbs and adpositions in a language where the nominal domain appears to be head-final and relatives also precede their associated nominal), whilst maintaining a single setting for the Head Parameter in that language. Analysis of V2 effects in West Germanic languages (Travis 1984) and of OV orders in Vata (Koopman 1984) can be seen to do essentially the same thing in relation to the behaviour of clausal XPs. In all cases, transformations, motivated by Case or other features, serve to interrupt a harmonic underlying head-initial or head-final word order.

While these approaches were highly constrained and empirically successful, they implied a notion of linear order at the narrow syntactic level, which came to be viewed as problematic by some. The reason for this was the increasing evidence that grammar is sensitive only to constituent structure and not to linear order. Thus children apparently fail to posit syntactic operations which are non-structure-dependent (Crain and Nakayama 1987), and modules such as Binding Theory seem to be sensitive only to hierarchical notions such as c-command and not to linear precedence. This was coupled with a renewed interest in something highlighted by Hawkins (1980, 1982): the observation that there are robust gaps in attested word orders. Kayne (1994) brings these various concerns to the fore, and proposes a

[17] Moreover, as Emonds (this volume) points out, the behaviour of specifiers, which might be considered the complements of phrasal projections, means that many so-called head-initial languages are, in a certain sense, really disharmonic (cf. also Hawkins 1980 on this).

theory whereby two asymmetric structural relations (dominance and asymmetric c-command) come together to determine linear order. The following section briefly considers this proposal and relates it to our principal concern: disharmonic word orders.

1.2.4 *Antisymmetry*

In a certain (controversial) sense, antisymmetry can be viewed as a return to the *grammaire générale* idea that there is a natural order of language (see section 1.1 above), though it is less clear in the generative paradigm that this is in any way connected to the order of thought. Kayne's (1994) Linear Correspondence Axiom (LCA) proposes the following direct connection between hierarchical structure and linear order:

(18) **Linear Correspondence Axiom** (Kayne 1994: 6)
[For a given phrase marker *P*, where *d* is the non-terminal to terminal dominance relation, *T* the set of terminals, and *A* the set of ordered pairs $<X_j, Y_j>$ such that for each j, X_j asymmetrically c-commands Y_j—TB/MS], d(*A*) is a linear ordering of *T*.

The LCA states that the dominance relation applied to the set of ordered pairs determined by asymmetric c-command relations gives a linear order of the set of terminals in a given phrase marker. Kayne argues, largely on an empirical basis, that the relevant linear order is precedence rather than subsequence so that the following holds:

(19) **Implication of the LCA**
A terminal X precedes a terminal Y iff a category dominating X asymmetrically c-commands a category dominating Y.

Given the further assumptions in Kayne 1994 (discussed by Kayne, this volume and Toyoshima, this volume), Kayne's proposal that (18) is a principle of grammar (UG) and that the relevant relation is universal precedence leads to what has become known as the Universal Base hypothesis. This refers to the fact that, in the absence of any movement, a phrase will have default spec–head–comp linear order, based on its inherent c-command relations.[18] It follows that all other surface word orders must be derived from this basic order via movement. This might, in a sense, be considered an extreme version of the proposals in Li, Travis, and Koopman (cited above in section 1. 2. 3), whereby disharmony arises via transformations from a harmonic base. In the case of the LCA, all orders diverging from consistent head-initial order must be movement-derived, even harmonic head finality. Importantly, though, independent

[18] This is assuming Kayne's (1994) category-based definition of c-command.

considerations (e.g. V-raising, the presence of a canonical subject position, the presence of a canonical topic position, and the activation of the left periphery more generally) entail that harmonically head-initial systems also require movement (compare standard generative analyses of English, French, Swedish, and Niuean[19] in this regard, for example). To the best of our knowledge, no language features only structures directly reflecting the Universal Base (see section 1.2.5 of this introduction, and Cinque, this volume, for further discussion).

One of the conceptual attractions of the LCA is that it apparently permits the eradication of linearity from Narrow Syntax (especially under Chomsky's 1995b reappraisal (p. 340)).[20] Kayne (this volume), however, claims that 'order' is still required at the narrow syntactic level, calling into question this apparent advantage. Its main empirical advantage, and the one that is of central importance here, is that it provides a potential explanation for several word-order asymmetries. Firstly, as Kayne notes, movement is very generally to the left in natural languages, and not the right. Rightward 'movement', where it occurs, has very different properties to its leftward counterpart, being subject to numerous restrictions (e.g. the Right Roof Constraint; cf. Ross 1967 for discussion). The LCA provides an immediate explanation for this fact, given Chomsky's (1993) Extension Condition: if movement is only possible to the root of the tree, and the root of the tree is the highest position in c-command terms (i.e. it c-commands all other nodes), then it follows that movement will always be leftwards (see Sheehan 2010 for an LCA-compatible account of extraposition). Kayne also notes other word-order asymmetries which similarly derive from the lack of rightward movement, such as the apparent lack of verb-penultimate languages and penultimate position effects more generally (cf. Kayne 1994 for further examples). A final example which is worthy of note concerns the apparent lack of *wh*-movement in OV languages. Kayne attributes this to the ban on multiple specifiers imposed by the LCA.[21] If head finality is derived via roll-up movement, then it follows that *wh*-movement to Spec-LCP will be banned where TP-to-spec-CP movement has taken place.

In its initial form, then, the LCA appears to be a restrictive theory of word order with the potential to account for a number of word-order gaps and asymmetries of the kind observed by typologists.[22] As Roberts (2007a: 13–14) notes, from the

[19] Cf. Pollock (1989), Holmberg and Platzack (1995), and Massam (2000, 2001, 2005) for discussion. Cf. also Alexiadou and Anagnostopoulou (2001, 2007) on the so-called *Subject in Situ Generalization* and its proposed universal consequences. For Chomsky (2013), who also specifically cites this Generalization, movement is very generally required in language systems, regardless of their head initiality or head finality, to facilitate labelling.

[20] Chomsky (1995a,b) takes the LCA to function as a linearization algorithm, applying only at the mapping to PF. See also Moro (2000).

[21] This stems from the category-based version of c-command which Kayne posits. In such a system, multiple specifiers of the same category mutually c-command each other and so cannot be linearized.

[22] The hypothesis, however, is not without its critics and we return to this matter in section 1.4.2 below.

perspective of the LCA, the notion of harmony reduces to a preference for either all or no heads to trigger comp-to-spec movement, a preference that may result from Input Generalization, the third-factor acquisition bias mentioned above. Mixed systems, by this standard, are then more difficult to acquire, and hence are predicted to be less frequent. Given Hawkins' and Dryer's findings that the majority of languages are actually disharmonic, however, it is not so clear that this is straightforwardly correct. Worth bearing in mind here, though, is undoubtedly Dryer's (1992a: 109, note 17) observation that 'the majority of inconsistencies among inconsistent [i.e. disharmonic—TB/MS] languages can be attributed to a small number of pairs of elements for which there is a skewed distribution, such as the general preferences for NRel order'. Also relevant is the extent to which individual disharmonic elements instantiate high-frequency items in the systems in question, as it is well known that high-frequency elements may exhibit irregular properties in relation to the more general system and it also seems to be the case that high-frequency irregulars are represented differently from regular forms (cf. i.a. Pinker and Prince 1991, 1994, and Marcus, Brinkmann, Clahsen, Wiese, and Pinker 1995 on the Dual Mechanism model).

1.2.5 *Remaining questions*

As already noted, the existence of harmony and of both harmonic and disharmonic orders creates particular difficulties for Principles and Parameters approaches to word-order typology, which predict that, all things being equal, any grammatical system must fall on one side or another of any cross-linguistic dichotomy. From this perspective, the fact that disharmonic languages are so prevalent is apparently positive for Kayne's LCA, which takes word order to be tied to language-specific movement operations. There is also a sense, though, in which the Universal Base Hypothesis is deeply surprising in the light of typological research. According to Greenberg (1963) and Hawkins (1980), the purest harmonic word-order types are VSO and SOV, with SVO being a mixed type with much less clear correlations. From Kayne's perspective, there is a sense in which SVO is the underlying order of all languages. Of course, whether this is a problem depends on what status is afforded to this universal base. If Cinque (this volume) is right, for example, then even head-initial surface orders are derived, in which case, there is clearly no sense in which SVO is predicted to be the most frequent order (contra Newmeyer 2005a). Nevertheless, as an anonymous reviewer observes, there is also no immediately evident sense in which a Kaynian approach, taken on its own, predicts the head initiality/head finality distribution facts thrown up by Dryer's (1989c, 1992a) genera[23]-based studies,

[23] In Dryer's (1989c) terms, a *genus* is a group of languages that are clearly closely related, with a time depth of 3500–4000 years. As indicated in Dryer's (2011e) genealogical language list contribution to the

in terms of which head-final systems consistently emerge as the most common cross-linguistically. Dryer (1989c) shows that in 111 genera (58% of the total at the time), OV order predominates,[24] giving a cross-linguistic tendency for objects to precede verbs, while Dryer (1992a) shows that postpositions are found in 119 genera out of 196, with the predicate preceding the copula in 76 out of 127 genera. Worth noting in relation to the distribution of head finality, however, is the constraint to which it appears to be subject, the Final-over-Final Constraint (FOFC), discussed in section 1.4 below, which Biberauer, Holmberg, and Roberts (2008a *et seq.*) have argued to follow from a version of Kayne's theory interacting with more general and, in part, possibly non-language-specific principles (see also the discussion in section 1.3.1 below of Hawkins' findings regarding constraints on the internal make-up of left-branching phrases).

What remains indisputable about the current theoretical situation, however, is that many questions remain unanswered. These include: Is there any evidence for the movement required to derive head-finality from a universal base? How are word-order generalizations to be captured, by movement or base-generation or a combination of the two? Given the attested variation, are word-order parameters to be stated for each (lexical/functional) category, for classes of categories, or for all categories subject to some defeasibility constraint? Is it then true that, in fact, anything goes, beyond, possibly, each category having to have a fixed internal order? If not, what generalizations can be made aside from the simple observation that most languages are tendentially head-initial or head-final? Is word order connected to other aspects of grammar, such as prosody? How stable are disharmonic systems and how are they acquired and thus preserved? Are there alternative linearization mechanisms which should be considered alongside the LCA and Head Parameter-based approaches? What role do extragrammatical factors play in the determination of word order, particularly in frequency-based terms? The papers in this volume aim to answer questions such as these, throwing new light on the nature of the relation between surface order and Narrow Syntax. In the following sections, we consider some of these issues in more detail.

World Atlas of Language Structures Online (*WALS Online*)—a list which distinguishes 510 genera and 212 language families, taking 2,678 languages into account—the choice of term is guided by the general idea of 'genus' in biological classification, where a genus is 'a set of species that are clearly closely related to one another'; thus his genealogical classification of languages is intended to be such that 'even a conservative "splitter" would accept [it]'.

[24] Considering the distribution of SOV, SVO, and VSO orders by genera, Dryer (1989c) obtains the following areal breakdown:

	Africa	**Eurasia**	**Aust-NewG**	**NAmer**	**SAmer**	**Total**
SOV	22	26	19	26	18	111
SVO	21	19	6	6	5	57
VSO	5	3	0	12	2	22

1.3 Word order and linguistic theory

1.3.1 *Frequency*

As noted repeatedly above, the preference for harmony across head–complement pairs is statistical rather than absolute (the question of whether absolute implicational universals also exist is separate and we leave it aside here; see i.a. Whitman 2008 and Biberauer, Holmberg, and Roberts in press for discussion). In recent years, there has been much debate as to the relevance of frequency skewings for theoretical linguistics. Hawkins (1980: 193) first proposed that 'the theory of Universal Grammar must include both implicational universals and universals of language distribution in the description and explanation of word order'. Hawkins objects to Greenberg's use of statistical universals where absolute implicational universals are empirically and theoretically superior, but he also notes that the preference for harmony is statistical rather than absolute. His point, then, is that while gaps can be attributed to UG principles, statistical trends must have a different kind of explanation. In his 1980 paper, he posits Cross-Categorial Harmony (CCH) to address the statistical nature of harmony in the word-order domain:

(20) **Cross-Categorial Harmony (CCH)**
The more similar the position of operands relative to their operators across different operand categories considered pairwise (verb in relation to adposition order, noun in relation to adposition order, verb in relation to noun order), the greater are the percentage numbers of exemplifying languages.
(Hawkins 1980: 98)[25]

[25] Hawkins thus draws a connection between markedness and frequency. This connection was also discussed by Lightfoot (1979: 77), who claimed that languages with more marked grammars will be less frequent than languages with less marked grammars, with diachronic changes being expected to involve changes from more to less marked (see also McCarthy and Prince 1994 on the Emergence of the Unmarked effects conceived of in Optimality-Theoretic terms, and Roberts 2007b for recent minimalistically oriented discussion). Evidently, statements of this type rest heavily on the interpretation assigned to the notion 'marked'.

A very different and, for a time, very influential approach to markedness considerations determining the cross-linguistic frequency of word-order patterns is found in Tomlin (1986). In terms of this approach, three interacting functionally motivated principles determine the observed frequencies: Theme First (more thematic information tends to precede less thematic information), Verb–Object Bonding (in a transitive clause, the object is more tightly 'bound' to the verb than it is to the subject—cf. also Baker 2009, 2010 for further discussion in a generative framework), and Animate First (Animate NPs tend to precede other NPs). Since SOV and SVO languages are consistent with all three of these principles, they are the most frequent, followed by VSO languages, which violate only Verb–Object Bonding, and then VOS and OVS languages, which violate the other two principles, but respect Verb–Object Bonding; since OSV languages violate every one of these principles, they are expected to be the least common. The difficulties with this type of functionalist markedness approach are well known (see Song 2012 for overview discussion and references), and, since it does not relate very directly to the question of (dis)harmony, we leave it aside here.

The reasoning here is that 'by reserving implicational statements for the task of distinguishing attested from non-attested co-occurrences, we can, therefore, formulate just one supplementary distributional regularity [i.e. CCH—TB/MS] . . . which captures generalizations which [statistical implications—TB/MS] are intrinsically unable to state' (Hawkins 1980: 232). This distributional regularity is arguably not a principle of UG, but rather a generalization over E-languages (see Chomsky (1986) for discussion of this term and how it differs from 'I-language').

Recently, this position has been endorsed by Newmeyer (2003, 2005b) and others (e.g. Whitman 2008). Newmeyer (2005b: ch. 3) has famously proposed that theoretical linguistics should concern itself only with possible and not probable languages. The frequency of a certain word order, he claims, is to be explained by external factors, such as patterns of diachronic change (as influenced by acquisition), functional pressures, or even arbitrary social pressures and language contact. Most specifically, he has endorsed Hawkins' (1994, 2004) processing-based account of the preference for harmony (cf. also Hawkins this volume). In this more recent work, Hawkins has moved towards an explicitly functional explanation for the CCH, based on his Performance–Grammar Correspondence Hypothesis (PGCH), which claims that languages have grammaticalized word orders which are efficient from a processing perspective. From this perspective, principles of processing efficiency such as Minimize Domains (MiD) (Hawkins 2004: 31, this volume) favour head adjacency as well as optional processes such as extraposition. Hawkins also argues that other more nuanced typological trends can be traced back to the PGCH: for example, the tendency for rigid VO languages to develop initial articles. This is because initial articles serve to construct NP and this is more efficient at the left edge in VO but not OV languages. A final asymmetry of this kind is that '[l]eft-branching phrases [. . .] are often more reduced and constrained in comparison with their right-branching counterparts [. . .]' (Hawkins this volume, p. 405). Thus prenominal relatives are often reduced compared to their postnominal equivalents.

The authors in this volume take different positions in this debate. Cinque's approach might lead us to expect that unmarked systems (i.e. those featuring fewer deviations from the 'ideal harmonic derivations' he postulates; see also discussion in section 1.3.2 below) should (all else being equal) be more frequent than more marked systems. This is not to say, however, that frequency cannot be affected by extragrammatical factors; hence his discussion of the relative scarcity of VOS orders in natural languages, despite the fact that the latter is one of the two abstract harmonic orders proposed by Cinque. In other papers, too, infrequent orders are taken to be ruled out by Universal Grammar or at least made difficult to generate by it. Thus, one of the reported benefits of Toyoshima's graph-theoretical linearization approach is that it 'accounts for the rarity of the other three logically possible [word-order] variations (VOS, OVS, OSV)' (Toyoshima, this volume: 360). Djamouri, Paul, and Whitman, on the other hand, argue explicitly against

the approach taken by Cinque whereby frequency is indicative of markedness. Instead, they attribute the rarity of head-final PPs in VO languages to diachronic factors. To the extent that diachrony is a reflex of acquisitional considerations, however, it could still be the case that markedness plays a role in determining the frequency of systems of different types.

Even if one accepts that third-factor pressures of the type alluded to in Chomsky (2005; see section 1.2.2 above) can interact with UG to give the Greenbergian correlations, it remains to be decided at what level this interaction takes place. Kiparsky (2008) seems to imply that the interaction operates at quite a deep level:

> The generative program opens up the possibility that [third factors—TB/MS] might have become biologized within UG itself… (Kiparsky 2008: 25)

Abels and Neeleman (2009) likewise speculate (as one possibility) that the requirement that a filler precede a gap might act 'as motivation for a grammatical principle stating that a moved constituent must be linearized at PF as preceding its sister' (cf. Ackema and Neeleman 2002 for the original discussion and motivation of these ideas). Once again, the implication appears to be that third-factor considerations could shape the very nature of UG. It is difficult to see how such a thing could be possible unless one adopts the evolutionary scenario put forth by Pinker and Bloom (1990), whereby UG evolved gradually via natural selection. The familiar objections to such an approach remain (see Fitch 2010 for overview discussion). A different perspective on the interaction of UG and third factors takes the latter to exert an influence only at the point of acquisition (parameter-setting in some models; cf. Biberauer, Holmberg, Roberts, and Sheehan 2010; Biberauer, Roberts, and Sheehan 2013 for more detailed discussion). The advantage to such a view is that third factors are not themselves 'biologized', but they rather serve only to constrain variation within a biologically determined variation space. If this speculation is on the right track, Newmeyer's (2005a: 105) conviction that '[t]he relationship between typological generalizations and I-language is therefore necessarily quite indirect' (cf. section 1.2.1 above) may not be so well-founded: our minds/brains may not *need* to 'know' about the typological status of components of mental grammars, as envisaged by Newmeyer, if typological facts can be shown to fall out as the consequence of the interaction between UG and suitably clearly formulated third-factor considerations.[26] As should be clear from this discussion, much work remains to be done to clarify the role of third factors in accounting for empirical skewings, and, more generally, to gain a better understanding of the significance or otherwise of frequency facts in the linguistic domain.

[26] As an anonymous reviewer points out, the question of how third-factor explanations relate to functional explanations of the kind proposed by Hawkins, Givón, Croft, Haspelmath, and others is just one of the issues that requires clarification. See Mobbs (2008, in progress) for partial discussion.

1.3.2 *On the nature of disharmony*

In relation to disharmonic word orders, Cinque opens the discussion by noting, as we have in this introduction, that extensive empirical research has undeniably reduced Greenberg's word-order correlations to statistical tendencies. He goes on to suggest that the search for *surface* word-order universals is arguably misplaced and that what Greenberg's generalizations really reveal is the order of Merge—the abstract universal hierarchical order of natural language.[27] The extent to which the linear order of categories in a given language departs from this abstract order can then be measured, yielding a finer-grained word-order typology. From this perspective the 'harmonic' orders are 'epiphenomenal', rather than basic, in that they represent cases in which the hierarchical order is systematically reflected at the linear level. Importantly, Cinque proposes that when word-order patterns are reconsidered from this perspective, a pervasive generalization nonetheless emerges:

(21) [W]hatever precedes the VP/NP reflects the order of Merge, and whatever follows is in the mirror-image of the order of Merge. (Cinque, p. 54)

The status of (21) is open to interpretation. If one assumes, as Cinque does, that surface word orders result directly from c-command relations, as proposed by Kayne (1994), and that there is a universal order of Merge, then word-order variation must be indicative of differences in movement operations between languages. As such, (21) can be taken to be a restriction on movement, as argued in Cinque (2005a, 2009a). Importantly, Abels (2007) and Abels and Neeleman (2009, 2012) suggest a different perspective: for them, processing considerations such as the filler–gap-related one discussed in the preceding section motivate leftward rather than rightward movement, but this consideration plays no role in determining base generation; consequently, c-command relations need not play any role in this domain, leaving open the possibility that First Merge could deliver both head-initial and head-final structures (cf. also Richards 2004 for a different argument for leaving open this possibility at the bottoms of projections, building on Epstein et al. 1998).

Other papers in the volume approach the topic of disharmony by considering a specific disharmonic language. Djamouri, Paul, and Whitman discuss evidence from Mandarin, which, they argue, represents a stable disharmonic system with both prepositions and postpositions, raising problems for the classical Head Parameter. In a careful empirical study, they show clearly that both prepositional and postpositional phrases can occur in postverbal argument position, but that there are restrictions governing the possibility of PPs in adjunct and subject positions: whereas certain kinds of postpositional phrases are banned from preverbal adjunct positions, prepositional phrases are systematically banned from subject position. Crucially, DPs

[27] What Greenberg (1963: 104) called the *proximity hierarchies*; cf. note 8.

are also banned from the subject position of existential sentences, suggesting that the distinction is not simply connected to the purported 'nominal status' of postpositional phrases (contra i.a. Li 1990; McCawley 1992; Huang, Li, and Li 2009). Inside DP, they show that, whereas prepositional phrases can modify only relational nouns, postpositional phrases can modify any kind of noun. These differences in distribution, they argue, serve as a cue to learners that postpositional phrases are pPs, whereas prepositional phrases are pure lexical projections, lacking a pP layer. pPs are analysed as being headed by a light preposition which attracts the complement of P to its specifier, yielding a surface head-final order. Prepositional phrases, on the other hand, are reduced PPs, without p, which therefore have a different distribution from full pPs. Strikingly, Djamouri et al. show that this system has remained reasonably stable from the first century BCE onwards. It must therefore be straightforwardly acquirable, potentially on the basis of specific distributional cues such as those mentioned above.

The same is arguably true in Basque, as Elordieta (this volume) discusses in a later section. Monolingual and bilingual Basque speakers acquire the disharmonic structures of Basque very early on without producing non-target-like orders at any stage of the acquisition process. The same has been shown to be true even in languages with complex V2 patterns (cf. Westergaard 2009a). More generally, it is well known that children acquire the word-order facts of their target language very early, at least as soon as they begin combining words (Bloom 1970; Brown 1973; and see Wexler 1998 on so-called *Very Early Parameter Setting* more generally). In fact, recent research suggests that prosodic cues have an effect on language-acquiring children at the *pre-lexical* stage already (cf. i.a. Christophe, Nespor, Guasti, and van Ooyen 2003; Bion, Höhle, and Schmitz 2007; May, Byers-Heinlein, Gervain, and Werker 2011). The theoretical implications of this fact are unclear, as Christophe et al. note. Elordieta (this volume), however, claims that the fact that Basque children acquire disharmonic structures as early as they do undermines the Universal Base Hypothesis in terms of which Kaynean Spec–Head–Comp structures are universal. She therefore interprets this acquisitional fact as support for the Head Parameter. Djamouri et al., on the other hand, claim that the stable disharmony in Chinese arises in a system which is underlyingly head-initial. Now, if head finality involves movement of a complement to some higher functional head, as i.a. Cinque, Djamouri et al., and de Vos propose in this volume, and children at the two-word utterance stage operate with truncated tree structures, lacking functional heads, as Rizzi (1993/1994) has suggested, then the data would appear to be problematic for the Universal Base Hypothesis. However, given that the truncation model is not uncontroversial (see i.a. Wexler 1998 for discussion; Guasti 2000 is a response) and that, according to Cinque, head-initial orders also involve the obligatory presence of functional heads, the data might rather be taken as evidence against that model of acquisition.

Cinque and Djamouri et al. focus on instances of disharmony where the same language contains some categories or even lexical items which are consistently

head-initial and others which are consistently head-final. A different kind of disharmony is observed in languages with variable word order, where one and the same category or lexical item either precedes or follows its complement, depending on context. Thus Cognola discusses Mócheno (German: *Fersentalerisch*), a Tyrolean variety spoken in the speech island *Valle dei Mócheni* (German: *Fersental*), in Northern Italy (Eastern Trentino). This variety displays mixed VO/OV orders, and Cognola focuses her discussion on the complex ways in which the VO/OV alternation interacts with the language's V2 property. Taking particular note of this latter feature of the Mócheno CP, Cognola argues that a similar V2 constraint regulates the structure of the VP, and that this is what is responsible for the observed VO/OV alternations. As such, wherever an XP is extracted via A-bar movement from VP, the lower V2 constraint forces the past participle to raise to a position high in the VP edge, triggering VO word order in an otherwise OV language. This analysis has important implications for our understanding of the V2 effect, and its potential explanation. It also provides empirical evidence for the existence of a VP periphery of the kind proposed by Belletti (2004).

1.3.3 *The connection to prosody*

Ancient rhetoricians such as Dionysius (60 BCE–7 CE) proposed that the order of words was determined by 'the rhythmic movement produced by the succession of long and short syllables' (Weil 1878 [1844]: 11–12). This position is not so far removed from recent claims in the generative literature that there is a close connection between prosody and word order. The original proposal for this connection came from Nespor and Vogel's (1982, 1986) Complement Law:

(22) **Complement Law** (Nespor and Vogel 1982)
Complements rather than heads are preferred locations for stress in all types of domains.

This has the following effect (as spelled out by Nespor, Guasti, and Christophe 1996; cf. also i.a. Cinque 1993).

(23) **Relative prominence in a prosodic phrase**
In languages whose syntactic trees are right-branching, the rightmost node of [a prosodic phrase—TB/MS] is labelled *strong*. In languages whose syntactic trees are left-branching, the leftmost node of [a prosodic phrase—TB/MS] is labelled *strong*. All sister nodes of *strong* nodes are labelled *weak*.

This idea is explored by Emonds (this volume) in relation to the syntax and morphology of French and English.

As Emonds points out, it has long been assumed that ordering in the morphological component proceeds on a different basis to ordering in syntax. Thus,

conventional wisdom has it that the Right-hand Head Rule applies in the morphology of many languages which are otherwise head-initial (Williams 1981; Hoeksema 1985, 1992; Scalise 1988; Lieber 1992). Emonds questions this widely-held belief, claiming that 'no head ordering statements pick out a domain coinciding with "Morphology" [in his terms—TB/MS]'. In particular, he takes issue with the commonly held claim that English is head-final in the morphological component and head-initial elsewhere. In the context of a Lieber (1992)-style approach to morphology, this is because a 'morphological component' cannot be meaningfully distinguished from a 'syntactic component'. Building on Lieber, Emonds' argument in this volume is that the properties that are typically ascribed to 'affixes' (and, thus, morphology) are such that it is impossible to distinguish 'morphological' elements from functional elements more generally. Specifically, affixes are typically said to (a) lack semantically interpretable features of the encyclopaedic type, beyond those which are syntactically active (cf. Chomsky 1965 for the original generative distinction between what Minimalists today call *semantic* and *formal* features), and (b) fail to contribute their own stress to word stress. Functional elements, according to Emonds, exhibit exactly the same properties. Against this background, he considers the headedness of English and French, both above and below what is traditionally taken to be 'the morphological level'. Emonds' central claim is that the universal default word order of natural language requires complements (and specifiers, which he views as a kind of complement) to precede heads. This is, in some respects, reminiscent of proposals by Haider, whose *Basic Branching Constraint* (BBC) also establishes head-final structures as the default option (cf. i.a. Haider 2000b for discussion and references). For Emonds, exceptions to the complement–head pattern are, however, possible, subject to UG principles and only in 'free' domains, defined as follows:

(24) **Free domain**
Domain Y is free if (i) no daughter of Y is an obligatorily bound morpheme and (ii) at least one daughter is an X^j that can further project, where X = N, V, A, P *and j = 0 or 1.*

This serves to force specifiers to be initial, as long as there is a single-specifier condition, so that no further projection of the phrasal head is possible. It also means that in the morphological component, only head-final structures are possible, as long as a bound morpheme is involved. Where morphemes are not bound, however, they can diverge from the universal default word order, as is the case in French. Thus Emonds shows that French compounds are head-initial only where they contain free morphemes, which can project further. He gives an extensive list of compounds of this type, including:

(25)

English right-headed compounds	**French left-headed compounds**
tanker **truck**	**camion** citerne
video **cassette**	**cassette** vidéo
bedroom **suburb**	**ville** dortoir

The relevant trigger for head–complement order in such domains is, of course, stress, which is uniformly on the right in French. This, combined with Nespor and Vogel's (1982) Complement Law (22), means that French is as left-headed as is permitted by UG. The reason why English is comparatively less left-headed stems from the fact that it exhibits left-hand stress, and so having left-hand heads would 'decrease compliance with the Complement Law'.[28]

While there is undeniable evidence for a connection between headedness and stress, as pointed out by Cinque's (1993) discussion of the Nuclear Stress Rule, it is not clear what the direction of causality is. Is it that the requirement for initial stress forces head-finality or is it rather that stress is sensitive to hierarchy rather than word order, and so tracks the position of complements? Emonds' contribution suggests that the former is true, and that, as Christophe et al. (2003) have proposed, stress is essentially an acquisition device which the child uses to set the word-order parameters of her language (arguably by the end of the first year of life[29]).

Hinterhölzl (this volume) proposes that the tendency towards harmony derives from a prosodic constraint of the following kind:

(26) **Mapping Condition to PF** (prosodic transparency)
A heavy syntactic constituent must appear on a dominant branch in prosodic phrasing if its containing phase is weight-sensitive. (Hinterhölzl this volume: 163)

This constraint ensures that, all else being equal, heavy (branching/recursive) constituents will align harmonically in a given language. In this system, disharmonic orders arise, then, (and can be preserved over time) where an optional and interfering constraint of the following kind applies:

(27) **Mapping Condition to LF** (scope transparency)
If *a* scopes over *b*, the Spell-Out copy of *a* should c-command the Spell-Out copy of *b*. (Hinterhölzl this volume: 163)

This less straightforward constraint appears to apply at the mapping to PF at the imposition of LF, in a way that is not possible within the standard Y-model. It forces

[28] Emonds discusses apparent counterexamples to this generalization from compounds involving prepositions.

[29] Worth noting here is that this statement should only be taken to refer to basic word-order parameters; the acquisition of discourse-sensitive word-order options like West Germanic scrambling is known to be delayed (cf. i.a. Schaeffer 1997, 2000 for discussion).

the object to be spelled out in its derived VP-external position in OV languages. Where it fails to apply, it follows that the lower VP-internal copy of the object is spelled out, yielding VO order in the context of the Kaynean model Hinterhölzl adopts, in compliance with (26). Importantly, this condition goes beyond the stipulation that German is OV whereas English is not, as it ties the OV order in German to the fact that it is scope-rigid, something which is apparently also true of other OV languages (cf. also Öztürk this volume).[30]

The applicability of (27) in German but not English also serves to explain the following contrast, first discussed by Haider (2000b): that (some) OV languages permit right-branching preverbal adverbials in apparent violation of the Head-Final Filter (HFF), whereas VO languages do not:

(28) a. John (more) often (* than Peter) read the book

b. Hans hat öfter (als der Peter) das Buch gelesen
Hans has more-often than the Peter the book read
'Hans read the book more often than Peter.' (Hinterhölzl this volume: 164)

In fact, German actually disallows extrapositon of these adverbials, something which is obligatory in English:

(29) a. John read the book more often than Peter

b. *Hans hat das Buch gelesen öfter (als Peter)
Hans has the book read more-often than Peter
(Hinterhölzl this volume: 164)

As Hinterhölzl notes, following Haider, this difference cannot be captured by a Head Parameter alone, as the relation between adverbial and *v*P is not one of complementation (cf. parallel difficulties for Head Parameter-based approaches already raised in relation to Dryer's 1992a non-correlation pairs in section 1.2.2 above). Adopting the Universal Base Hypothesis, Hinterhölzl proposes that the differences between VO and OV languages of this kind stem not from the fact that one involves movement whereas the other does not, but rather from the different sizes of the moved constituents, as motivated by the differing 'mapping conditions' (26) and (27) in the two kinds of languages (cf. also Cinque 2004, this volume for a similar proposal).[31]

[30] The VO/OV word-order variation attested in Old High German, he claims, stems from the at the time stronger force of the condition on scope transparency in conjunction with the weaker force of the condition on prosodic transparency. This OT-like model can thus potentially offer a prosodic explanation for harmony, a description of synchronic disharmonic systems and an account of diachronic word-order change.

[31] Hinterhölzl argues that movement of the *v*P to a pre-adverbial position (which he terms *intraposition*) is semantically motivated by the need for the adverbial to become a predicate of *v*P. As the moved vP functions effectively as a subject after movement, this is A-movement. The ability of *v*P to function first as a predicate and then as an argument is linked to the phase-based model which Hinterhölzl adopts.

In a VO language, *v*P intraposition will give rise to an optimal prosodic unit where the adverbial, which is heavy (in his terms), occupies a right branch (in his terms) and receives stress. This also favours pied-piping of the PP if subsequent vP movement is required (e.g. in the presence of another adverbial modifier), correctly yielding the following (unmarked) word order:

(30) $[_{IP}$ John$_i$ $[[[_{vP}$ t_i visited them$]$ $_k$ in Vienna t $_k]_j$ on Friday $t_j]]$

In German, on the other hand, event-related (time, manner, and place) adjuncts can either precede or follow the verb, with the difference crucially correlating with a scope-related interpretive difference (see (31)). Significantly, in both languages, adverbials surface in the unmarked preverbal order T>P>M, with M>P>T being the unmarked order postverbally.

(31) a. weil Hans oft im Kaffeehaus sitzt
since Hans often in-the coffee-house sits
'as Hans often sits in the coffee house'

b. weil Hans oft sitzt im Kaffeehaus
since Hans often sits in-the coffee-house
'as Hans, when he is in the coffee house, often sits'

On Hinterhölzl's proposal, this is explained because 'placement of adjuncts is weight-insensitive' in German (p. 181) because of (27): adverbials scope over *v*P and so must precede them, even if this violates (20). If further adverbials are present, then it follows that pied-piping of the adverbial will be optional.

Tokizaki and Kuwana (this volume), in turn, attempt to provide a different explanation for the connection between stress and word order. They first observe that the juncture between terminals in right-branching structures is longer than the juncture between those in left-branching structures. They take this as evidence that left-branching structures are actually compounds (an idea which Zwart 2009b has also explored). The repercussion of this is that stress patterns in left-branching structures must adhere to the word-stress pattern of a language. Right-branching structures, on the other hand, need not mirror the word-stress rules of a given language. This fact, it is argued, serves to explain the close connection between word stress and word order first noted by Nespor and Vogel (1982) and also explored by Emonds (this volume). Although left-branching structures are 'compounds', they nonetheless receive their stress derivationally, rather than lexically, presumably via the Nuclear Stress Rule (this idea is, of course, also compatible with Lexical Morphology models such as that of Halle and Mohanan 1985, which assign compounds to a third stratum of stress assignment, following irregular inflection and derivation, and regular derivation respectively). This means that stress falls on the most embedded constituent in a given structure. In the case of VO/OV orders, they claim, O is always

more embedded than V, as it is potentially branching (and this is true even where it has moved to Spec-VP). In a language with initial stress, it follows that OV order will be permitted and arguably required, as this movement serves to re-establish initial stress at the level of the compound. Where a language has right-edge stress, on the other hand, movement is blocked, as it would move the stress too far to the left, violating the word-stress rules of the language. Matters are more complex with other stress patterns, but Tokizaki and Kuwana nonetheless claim that the attested word-order patterns are roughly as predicted by the proposal.

It is interesting to note that while these three approaches are similar in spirit, they differ substantially as to the relationship they posit between grammar and linear order. Hinterhölzl, and Tokizaki and Kuwana both adopt a version of the LCA whereby head finality is derived via movement (albeit of different kinds). Emonds, on the other hand, argues against the LCA and proposes an alternative whereby SOV is the universal base (see again Haider 2002 for another OV base-order proposal). That proponents of both types of proposal are able to appeal to (22) reinforces the fact noted by Christophe et al. (2003) that Nespor and Vogel's Complement Law is consistent with both movement-derived and base-generated approaches to head finality.

In the following section, we consider movement-derived head finality in more detail.

1.4 The question of Antisymmetry

1.4.1 *Head-Complement order, movement, and the derivation of OV languages*

Whether one accepts the LCA in its strongest form or not, it seems reasonably clear that the surface word order in many languages is derived via movement. OV languages, for example, come in at least three guises:[32]

(32) i. DP-V-X (Nupe, Mande (Niger-Congo) (postpositional) and Päri (Nilo-Saharan, Nilotic, Sudan), Tobelo (West Papuan, North Halmaheran, Indonesia), Iraqw (Afro-Asiatic, Southern Cushitic, Tanzania), and Neo-Aramaic (Afro-Asiatic, Semitic, Israel) (prepositional))

[32] Worth noting here is that the OV languages discussed here can, in generative terms, be thought of as differing in relation to the extent to which they are head-final in the clausal context: type I is only minimally head-final, while type III is maximally head-final, with type II occupying quite a broad spectrum in between. Not discussed here, but also relevant to more fine-grained consideration of the typological question we raise here is the matter of the headedness of non-clausal categories (e.g. nominal and adpositional heads).

ii. DP/PP-V-CP (West Germanic, Turkish, Persian, Hindi, plus Lokaa and Vata with added complications[33])
iii. rigid OV (Japanese, Malayalam, Sinhala, Korean, Kannada)

Movement-based approaches to head finality are highly plausible for Type I languages, where only DP direct objects surface in a preverbal position (cf. Roberts 2007a: 14 for discussion). Kandybowicz and Baker (2003) (henceforth K&B) discuss one relevant language, Nupe, which displays this word order in perfect clauses. In all cases, only a single argument can and must precede V:

(33) Musa á etsu yà èwò.
Musa PRF chief give garment
'Musa has given the chief a shirt.'

Manner adverbs can follow V (but cf. Mous 1993 on Iraqw, Cushitic, another Type I language):

(34) Musa á nakàn ba sanyin.
Musa PRF meat cut quietly
'Musa has cut the meat quietly.' [Nupe, K&B (2003: 123)]

Given that DP direct objects, unlike PP/CP arguments and adverbials, enter into an Agree relation with a higher functional head (v in Chomsky 1995a), which assigns them Accusative Case, it seems highly plausible that OV order in this language (and others like it) is derived via A-movement, as K&B propose.

In Type II languages, it is less immediately clear that the movement-based approach has any immediate empirical advantages over a base-generation account, but advantages arguably emerge upon closer consideration, given certain assumptions. In Dutch, which might be considered a canonical Type II language, DPs, predicative PPs and APs, and non-sentential adverbs *must* precede V. CP arguments *must* and certain (phrasal and sentential) adverbials *can* also follow V (cf. Zwart 1997a,b on Dutch, and Baker 2005 on Lokaa, which is SOVCP only in negative clauses). The traditional account of this pattern takes these languages to be base-generated OV languages (cf. i.a. Koster 1975 and den Besten 1977/1983 for Germanic, Kural 1997 for Turkish, and Karimi 2005 for Persian). Stowell's 'Case Resistance Principle' (stating that Case may not be assigned to a category bearing a Case-assigning feature, Stowell 1981: 146) then forces CPs to be extraposed to avoid Case assignment. Certain problems arise for such an account, however, notably the fact that it seems to wrongly predict that PPs should pattern with CP rather than DPs. A further problem stems from the fact that extraposition should move the CPs to a non-argument position, predicting they will be strong islands, but they actually appear to permit subextraction in a number of Type II languages (cf. i.a. Zwart

[33] See Baker (2005) for discussion.

1997a,b on Dutch, Mahajan 1990 on Hindi, Aghaei 2006 on Persian, and Biberauer and Sheehan 2012 for overview discussion):

(35) Hoe heeft Piet gezegd dat Jan zich t gedragen heeft?
how has Piet said that Jan himself behaved has
'How did Piet say that Jan behaved himself?' [Dutch Zwart 1997a: 66]

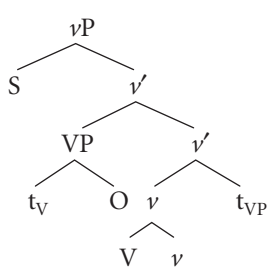
Zwart (1997a,b) proposes a movement-based analysis of Dutch, which accounts for these facts and might plausibly extend to all Type II languages. DP direct objects raise to a Case position above V, giving DP–V order. PP/AP predicates, however, move to Spec-PredP, a distinct preverbal position, while CPs simply remain *in situ*. Zwart provides evidence that material can intervene between V and a small clause complement, strongly suggesting that the latter at least *can* move (Zwart 1997a: 103):

(36) De kwast waar Jan de deur rood [$_{PP}$ mee t] verft
the brush where Jan the door red with paints
'The brush that Jan paints the door red with'

More generally, the Germanic languages appear to conform to a 'size'-based generalization, which can, following Wurmbrand (2001: 294), be stated as 'the "bigger" a complement . . . , the more likely it is to extrapose; the "smaller" the complement . . . , the more likely it is to occur in intraposed position' (cf. also i.a. Hinterhölzl 2006: 15). Biberauer and Roberts (2008) relate this pattern to the Final-over-Final Constraint (see section 1.4.4 below), predicting its occurrence to extend beyond Germanic (see also Biberauer and Sheehan 2012). For present purposes, the important point is that structures like (36) are, like those involving nominal and adpositional complements, clearly amenable to a leftward movement analysis of the sort an antisymmetric analysis would lead us to expect.

An alternative movement-based account of Type II languages is provided by i.a. Hinterhölzl (2006, this volume) for German, Haegeman (1998) for West Flemish, Baker (2005) for Lokaa, Biberauer (2003) for Afrikaans, and Biberauer and Roberts (2008) for West Germanic generally. These approaches basically derive head finality via X-movement followed by remnant XP movement, which serves to carry X's complement to a pre-head position (cf. Roberts 2007a: 15–16 for a simple overview):

(37)



This captures the fact that in Type II languages, everything except CP (and some adverbials) precedes V, even elements that cannot move independently (e.g. particles). It also avoids the positing of distinct functional projections to house DP and predicate arguments. The question remains, though, why CP fails to move to a preverbal position with the VP-remnant. In his discussion of Lokaa, Baker (2005) simply proposes that CPs are extraposed before VP movement. Biberauer and Roberts (2008), on the other hand, propose that CPs are radically spelled out when the vP phase is complete: this entails that the CP is removed from the syntactic computation prior to vP movement, being linearized immediately, with the result that it is, as the first constituent to be sent to the interfaces, spelled out in final position. While these accounts are not implausible, the fact that postverbal CPs remain transparent for subextraction remains potentially problematic, to the extent that it holds in all Type II languages.[34] The remaining problem is that, from a movement perspective, there is no deep explanation why PPs/APs pattern with DP in being preverbal, whereas CPs do not. From a remnant movement perspective, there is no clear explanation why CP remains postverbal, though proposals exist. The pattern might, for example, potentially be explained by the Final-over-Final Constraint/FOFC (in some formulation) if DP/PP/AP are not subject to FOFC (at least in some languages), whereas CP *is* (cf. Biberauer and Sheehan 2012 for an account along these lines and see section 1.4.4 below for discussion of FOFC).

In the case of Type III languages, the evidence for movement is somewhat scarcer. In rigid OV languages, all arguments and adverbials precede V. While a remnant movement account of the kind proposed by Biberauer and Roberts (2008) and Baker (2005) can clearly replicate this fact, it is not clear that there is any evidence in favour of such movement. These languages remain, then, from an Antisymmetry perspective, the most controversial in status, and the arguments for deriving their surface orders via movement are largely conceptual, stemming from the desire for a uniform approach to word order. Öztürk (this volume) discusses a number of Turkic languages, which might be classed as Type III languages, in relation to the LCA, Similarly, Elordieta (this volume) discusses Basque, a language which is predominantly head-final, but which displays some degree of variability and disharmony.

More generally, a number of the papers in this volume contribute to the ongoing debate surrounding Antisymmetry. Kayne (this volume) can be considered a reassertion of the LCA in its strongest form, rejecting Chomsky's (1995a) influential reappraisal of this axiom as a linearization algorithm. Barrie (this volume) further shows that it is possible to analyse the OVS language Hixkaryana in LCA-compatible terms. In earlier sections, Cognola proposes an LCA-compatible account of OV/VO alternations in Mochenò, and Hinterhölzl of basic word order in German. The other

[34] If Huang's (1982) Condition on Extraction Domain is correct, the possibility of subextraction necessarily means that the postverbal CPs under discussion here were base-generated in that position.

papers in the section under discussion here are, however, more critical of the LCA: Öztürk and Elordieta question the validity of the LCA for Turkic languages and Basque respectively. In addition, Toyoshima, who offers a novel alternative to the LCA (see section 1.4.3 below), raises conceptual objections to the version of the LCA adopted in many discussions of word order. He notes that the SVO 'universal base' is largely an artefact of (i) the definition of c-command adopted by Kayne and (ii) the arbitrary choice of precedence over subsequence, which is a primitive of the axiom. These challenges are acknowledged and addressed by Kayne in his contribution.

Kayne's chapter can be divided into three distinct parts. Firstly, he reviews the evidence that OV order must be derived in many languages, drawing on a wealth of research since Kayne (1994). For this reason, there must be more to word order than the Head Parameter. Secondly, he reviews further empirical evidence that syntax is not symmetrical, drawing on work from generative and typological studies: the predominance of left-dislocation over right-dislocation, the fact that clitics often surface further to the left than full DPs, the fact that agreement on X is often suspended where DP follows X, but not where DP precedes X, certain asymmetries in relative clause formation, the fact that serial verbs surface in the same order in so-called head-initial and head-final languages (Carstens 2002, also discussed in Kayne 2003a), and the fact, observed by Zwart (2009a), that coordination always requires a coordinator between the two coordinates, which follows if the coordinator is always a head and the basic linear order of a coordinate phrase is Spec–Head–Comp. The final asymmetry discussed is of a different order to the others and indeed to those discussed previously by Kayne (1994, 2003a), though. Kayne notes that while some languages disallow backwards pronominalization (38), no language is known to block forwards pronominalization (39):

(38) The fact that he's here means that John is well again.

(39) The fact that John is here means that he's well again.

This asymmetry, he notes, cannot be stated purely in terms of c-command, as the R-expression and pronominal enter into no c-command relation in either (38) or (39). This cross-linguistic asymmetry Kayne, then, takes to suggest that precedence itself is syntactically encoded, contra Chomsky (1995a,b) and much subsequent work on the LCA.

The remainder of his chapter addresses the more fundamental question of why natural language is antisymmetric. The argument can be deconstructed roughly as follows:

(40) a. Probe–Goal search shares the directionality of parsing and of production.
b. Production proceeds from left to right.
c. Therefore, Probe–Goal search proceeds from left to right.
d. Probes are heads and Goals are contained in their complement domain.
e. Therefore, heads must precede complements.

In essence, this means that 'FL has incorporated an abstract counterpart of temporality' (Kayne this volume: 234). The validity of this argument hinges on the acceptance of (40a), which Kayne proposes without any specific justification. If Probe–Goal directionality is merely stated in hierarchical terms, then it is consistent with either Comp>Head or Head>Comp linear order (at least for non-head complements, i.e. those that do not fall foul of the much discussed 'bottommost pair' problem—see i.a. Chomsky 1995b: 418; Richards 2004 for discussion). In a sense, the argument in (40a–e) is a means of reducing (e) to (a), but it is by no means a justification of (e). The discussion of Specifier>Head order is more opaque. Here Kayne rejects Abels and Neeleman's (2009, 2012) claim that leftwards movement be elevated to the status of an axiom, and maintains the idea that it should derive from the more general fact that specifiers precede heads. His proposal is that specifiers merge with heads directly, rather than phrases (cf. Sheehan in press for a similar proposal). The problem is that even in instances of internal Merge, the head probes the (derived) specifier prior to movement under standard assumptions, so that the predicted order is actually Head>Specifier according to Kayne's assumptions. Kayne denies this probing relation, the implication being apparently that Specifier probes Head. The fact that pair-Merge must lead to immediate precedence is also invoked, so that the only way for H to pair-Merge with two distinct phrases is if they surface on different sides of H. The need for immediate precedence also serves to rule out multiple specifiers (given that the merger of two phrases is banned). Note, though, that, as was the case in previous versions of antisymmetry theory, a non-projecting specifier (head) should still be able to merge with a phrase without projecting, giving rise to multiple non-branching specifiers (cf. Guimarães 2000). Ultimately, then, the fact that specifiers must precede heads also follows from (a). If (a) is rejected, it follows that neither conclusion can be maintained. At the end of his contribution, Kayne makes the radical proposal that the Specifier/Complement distinction may not be reflected structurally, so that branching is ternary rather than binary. Among other things, this has the effect that binding theory can no longer be stated in terms of c-command (unless the definition of c-command is amended).

Barrie's contribution considers the properties of Hixkaryana, an OVS language. He proposes a smuggling analysis (cf. Collins 2005) whereby the object is carried past the subject via VP-fronting. This serves to derive the fact that OVS languages appear to disallow scrambling, as the object must remain low enough to be carried along via VP-fronting. He contends that this correlation provides empirical support for the LCA, as Head Parameter-based analyses of OVS orders fail to capture the correlation of these two properties. As an anonymous reviewer pointed out, though, this is not necessarily the case as, however OV order is derived, a ban on fronting anything larger than VP will also serve to rule out OXVS order. The generalization, then, while it provides strong evidence that OVS order is derived via VP movement, for the reason Barrie outlines, does not necessarily bear on the derivation of OV order. In fact, given

anti-locality (cf. Abels 2003), it could even be argued that the ban on OXVS order is an empirical challenge for his analysis. In the same way that the general availability of scrambling in SOV languages is taken as crucial evidence that OV is a derived order, it could also be concluded that the ban on OXVS is evidence that OV order is *not* derived via movement. This issue is partially addressed by Barrie via his discussion of Japanese vs Germanic OV. He claims that in the former case, a ghost or proxy (cf. Nash and Rouveret 2002) projection gives rise to rigid OV order, irrespective of Case, whereas in the latter, OV order results from (presumably Case-related) object shift. An alternative interpretation (consistent with the retention of a more restricted Head Parameter) would be that Germanic OV languages derive OV order via movement, whereas Japanese base-generates it.[35] Evidently, then, there are still a great many open questions relating to the structural analysis of different types of OV order, and the role that Antisymmetry has to play in increasing our understanding of it.

1.4.2 *Problems with the LCA*

The chapters by Öztürk and Elordieta argue that the LCA is not suited to enhancing our understanding of languages exhibiting a considerable amount of head finality. Öztürk's contribution focuses on some empirical challenges from the syntax and semantics of postverbal constituents (PVCs) in two lesser-studied OV languages, Khalkha and Uyghur. Building on Kural's (1997) challenge to the LCA, Öztürk claims that postverbal constituents in Turkish and Uyghur pose recalcitrant difficulties for the idea that asymmetric c-command maps to precedence. She argues at length that PVCs in these languages display the properties of movement, displaying island sensitivity of the same kind as constituents that have undergone leftward movement and failing to co-occur with resumptive pronouns. Importantly, this is different from what is observed in other OV Altaic languages such as Japanese and Khalkha, where PVCs are clearly base-generated. Any attempt to account for the Uyghur facts via leftwards movement, she argues, will prove problematic, as these languages are otherwise scope-rigid, meaning that scope tracks surface c-command relations (cf. Hinterhölzl's (27), discussed above). As PVCs can take either wide or narrow scope

[35] As Biberauer (2008b) notes, third-factor considerations—specifically, representational economy at the level of the child's stored grammar (I-language)—could plausibly lead the child to 'reanalyse' a grammar for which all heads are associated with a linearization-related movement diacritic (*EPP-feature* in Barrie's terms) as one in which all heads simply take their complements to the left, i.e. in which these heads are diacriticless and PF imposes across-the-board head finality (cf. Richards 2004, 2009 for discussion of the mechanics of a PF Head Parameter). What seems crucial to approaches seeking to combine antisymmetric derivations of head-final orders with Head Parameter-based ones is that there be a principled basis on which we (and language acquirers) are able to determine which of the available options (movement or 'base generation'/a PF parameter) underlies input from different types of head-final languages. In the absence of a principled distinction of this type, the acquisition task arguably becomes intractable, while the syntactician's task of unambiguously characterizing what underlies the head finality of specific systems is also compromised.

with respect to preverbal quantifiers, rightward movement would seem to be required in order to capture these facts straightforwardly. Öztürk claims that, in the absence of rightward movement, the facts elude explanation, forcing one to abandon scope rigidity and to posit unmotivated ad hoc movements. According to Öztürk, the reason why Turkish/Uyghur pattern differently from Japanese/Khalkha (and other familiar languages such as English, German, French, etc.) is that the former, but not the latter, have an EPP requirement. She argues that the presence of a left-hand specifier blocks the possibility of a right-hand specifier, and hence right-ward movement. To ensure that rightward movement is not possible at any structural level (vP, TP, CP, etc.), the proposal seems to require that all clausal heads bear an EPP feature: this is what is necessary to trigger the projection of a leftward specifier, which then suppresses the projection of a rightward specifier. It is not, however, clear that this prediction holds up. The correlation between a lack of subject-related EPP (i.e. an EPP feature on T) and rightward movement which Öztürk's paper highlights seems to hold in a number of languages, but why this should be the case remains opaque.

The discussion of Basque in Elordieta (this volume) also touches on this subject, as Basque is another language which displays both VO and OV orders, depending on context. In her contribution, Elordieta argues against Haddican's (2004) anti-symmetric approach to this alternation. Her main claim is that a non-antisymmetric (Head Parameter-based) account of the facts is possible, so that the Basque OV/VO facts cannot be interpreted as specifically lending support to the Universal Base Hypothesis, contrary to what has previously been claimed. In fact, what emerges from Elordieta's discussion is that it is very difficult to construct robust empirical evidence against either approach in the Basque case: without a principled basis for distinguishing between movement- vs Head Parameter-imposed OV and VO orders and without constraints on what may move where, both approaches can derive the attested orders (cf. note 35 and also the discussion in Abels 2007). From an acquisitional perspective, we might expect the child to opt for simpler rather than more complex representations of the available input, an expectation which might initially seem to favour non-antisymmetric approaches. This is true, though, only if such simple representations are made possible by UG. Among other considerations, such approaches, however, appear to offer no insight into cross-linguistically unattested word orders (such as those ruled out by the Final-over-Final Constraint, which we turn to in section 1.4.4, for example), and it is also not so clear that a Head Parameter-only approach to the types of OV languages discussed in section 1.4.1 will facilitate insight into this instance of attested variation. Taking this into account, it becomes clear that the (in some cases) greater representational simplicity associated with Head Parameter-oriented interpretations of OV/VO word-order variation cannot be taken, in isolation, to signify the superiority of approaches of this type.

1.4.3 *Novel approaches to disharmonic word order*

Aside from papers highlighting difficulties with Head Parameter- and Antisymmetry-based analyses, the volume also contains two papers proposing novel analyses of disharmonic word-order phenomena. De Vos (this volume) discusses the various kinds of adpositional phrases attested in Afrikaans and Dutch. Building on previous work by Oosthuizen (2000), Biberauer and Folli (2004), and Biberauer (2008a), he shows that, whereas prepositional phrases in these languages receive a locative interpretation, postpositional and circumpositional phrases are always directional, encoding directed motion:

(41) Disharmonic word orders in the Afrikaans adpositional domain

a. Ek loop **in** die kamer
I walk in the room
'I walk around inside the room.' [head-initial adposition]

b. Ek loop die kamer **in**
I walk the room in
'I walk into the room.' [head-final adposition]

c. Ek loop **in** die kamer **in**
I walk in the room in
'I walk into the room.' [circumpositional adposition]

He further argues that postpositions, unlike prepositions, Agree with the DP which they select. It is this Agreement dependency, he argues, which gives rise to either (i) postpositions or (ii) circumpositions. There is a sense in which de Vos's proposal appears to be a notational variant of existing approaches to movement. Landau (2007), for example, proposes that the EPP feature is essentially the need to realize a certain narrow syntactic Agree dependency overtly at PF (via displacement of the Goal), and more generally, it is typically assumed that A-movement at least is a PF effect which is parasitic on existing Agreement dependencies. De Vos's account, however, differs from existing approaches in that it proposes that *circumpositional* structures fall out from the same PF requirement. Essentially, faced with the paradoxical need to realize both P→DP (selection) and DP→P (Agree) dependencies linearly, the PF component has two options: (i) DP–P–DP or (ii) P–DP–P. If option (i) is chosen, a language-specific requirement forces deletion of the rightmost DP, yielding a postpositional phrase.[36] If option (ii) is chosen, then in Afrikaans, no chain reduction is required and both copies of P are retained, yielding a circumpositional phrase. This approach provides an elegant PF account of the variation in

[36] Exactly why or how this linear deletion operation applies remains unclear, however.

Afrikaans, which makes no allusion to the Head Parameter. Despite its elegance, however, it raises certain questions, the most obvious being: given the pervasive existence of agreement dependencies in natural language; for example, why is it that doubling is not more common?

Toyoshima develops a graph-theoretic linearization proposal originally made by Kural (2005). In terms of this proposal, the major word orders, defined in relation to S, O, and V, all result from PF undertaking different tree traversals of the same underlying structures. In particular, Toyoshima's version of Kural's algorithm derives the three commonest word-order variants (SOV, SVO, and VSO) from a single structure. Furthermore, it is also argued to account for the rarity of the other three logically possible variations (VOS, OVS, OSV) and for disharmonic word-order patterns of the type found in Vata and German. The general approach addresses some of the challenges facing Kural's approach, notably its reliance on notions such as 'right' and 'left' at the narrow syntactic level, which Toyoshima replaces with the notions 'consanguineous' (i.e. dominated node of which the label is non-distinct from that of the parent node) and 'adopted' (i.e. dominated node of which the label is distinct from that of the parent node). Challenges to this approach would seem to include its ability to capture *wh-* and head movement. Of greatest relevance to the concerns in this volume, however, is what needs to be said about disharmonic word orders. In the German case, for example, it is necessary to stipulate a 'parametric feature' on phase heads that is visible at PF, specifying the order of traversal for specific phrases. Additionally, like Head Parameter- and Antisymmetry-based approaches to disharmonic word orders, the traversal approach, however, seems to fall short when it comes to being able to account for a striking gap in the attestation of such word orders: those ruled out by the Final-over-Final Constraint, to which we now turn.

1.4.4 *The Final-over-Final Constraint*

Recently, it has been pointed out in several places that the empirical evidence given in favour of the LCA is incomplete. While Kayne and others have provided strong empirical support for the lack of right-hand specifiers, it is claimed that the lack of lefth-hand complements is less well evidenced (cf. Richards 2004, 2009; Abels and Neeleman 2009, 2012). Some of the data discussed in this volume provide potential evidence of the required kind. Firstly, there is Barrie's claim that *OXVS order is banned, for principled reasons. Unfortunately, it is not clear (i) that this bears on the lack of a Head Parameter or (ii) that this follows necessarily from Antisymmetry. Even if OV order were derived via a Head Parameter, it would be sufficient to ban movement of any constituent larger than VP in order to rule out OXVS, as long as specifiers are always to the left so that OVS order is necessarily derived via movement. The crucial evidence that the order of heads and complements is regulated in