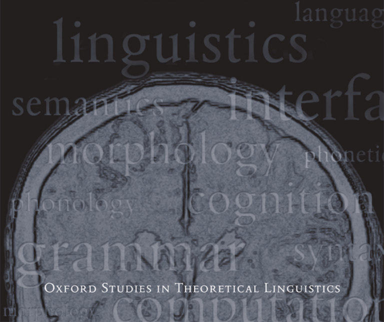
OXFORD
LINGUISTICS

Causation in Grammatical Structures

edited by

BRIDGET COPLEY AND
FABIENNE MARTIN



OXFORD STUDIES IN THEORETICAL LINGUISTICS

Causation in Grammatical Structures

OXFORD STUDIES IN THEORETICAL LINGUISTICS

Recent titles

33 Events, Phrases, and Questions
by Robert Truswell

34 Dissolving Binding Theory
by Johan Rooryck and Guido Vanden Wyngaerd

35 The Logic of Pronominal Resumption
by Ash Asudeh

36 Modals and Conditionals
by Angelika Kratzer

37 The Theta System
Argument Structure at the Interface
edited by Martin Everaert, Marijana Marelj, and Tal Siloni

38 Sluicing
Cross-Linguistic Perspectives
edited by Jason Merchant and Andrew Simpson

39 Telicity, Change, and State
A Cross-Categorial View of Event Structure
edited by Violeta Demonte and Louise McNally

40 Ways of Structure Building
edited by Myriam Uribe-Etxebarria and Vidal Valmala

41 The Morphology and Phonology of Exponence
edited by Jochen Trommer

42 Count and Mass Across Languages
edited by Diane Massam

43 Genericity
edited by Alda Mari, Claire Beyssade, and Fabio Del Prete

44 Strategies of Quantification
edited by Kook-Hee Gil, Steve Harlow, and George Tsoulas

45 Nonverbal Predication
Copular Sentences at the Syntax-Semantics Interface
by Isabelle Roy

46 Diagnosing Syntax
edited by Lisa Lai-Shen Cheng and Norbert Corver

47 Pseudogapping and Ellipsis
by Kirsten Gengel

48 Syntax and its Limits
edited by Raffaella Folli, Christina Sevdali, and Robert Truswell

49 Phrase Structure and Argument Structure
A Case Study of the Syntax-Semantics Interface
by Terje Lohndal

50 Edges in Syntax
Scrambling and Cyclic Linearization
by Heejeong Ko

51 The Syntax of Roots and the Roots of Syntax
edited by Artemis Alexiadou, Hagit Borer, and Florian Schäfer

52 Causation in Grammatical Structures
edited by Bridget Copley and Fabienne Martin

For a complete list of titles published and in preparation for the series, see pp. 452–3.

Causation in Grammatical Structures

Edited by

BRIDGET COPLEY AND

FABIENNE MARTIN

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 Edition published in 2014

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: 2014936065

ISBN 978–0–19–967207–3 (hbk.)
ISBN 978–0–19–967208–0 (pbk.)

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

Contents

Part II. From Causal Meanings to Causal Structures

General preface

The theoretical focus of this series is on the interfaces between subcomponents of the human grammatical system and the closely related area of the interfaces between the different subdisciplines of linguistics. The notion of 'interface' has become central in grammatical theory (for instance, in Chomsky's Minimalist Program) and in linguistic practice: work on the interfaces between syntax and semantics, syntax and morphology, phonology and phonetics, etc. has led to a deeper understanding of particular linguistic phenomena and of the architecture of the linguistic component of the mind/brain.

The series covers interfaces between core components of grammar, including syntax/morphology, syntax/semantics, syntax/phonology, syntax/pragmatics, morphology/phonology, phonology/phonetics, phonetics/speech processing, semantics/pragmatics, and intonation/discourse structure, as well as issues in the way that the systems of grammar involving these interface areas are acquired and deployed in use (including language acquisition, language dysfunction, and language processing). It demonstrates, we hope, that proper understandings of particular linguistic phenomena, languages, language groups, or interlanguage variations all require reference to interfaces.

The series is open to work by linguists of all theoretical persuasions and schools of thought. A main requirement is that authors should write so as to be understood by colleagues in related subfields of linguistics and by scholars in cognate disciplines.

This volume brings together cross-disciplinary perspectives on the notion of causation, linking psychological, philosophical, and linguistic perspectives together in an attempt to understand the notion of causation in a more profound way. The volume extends traditional linguistic work on how causation is realised in a wider range of understudied languages and links it to philosophical and cognitive perspectives on how we should understand the notion of cause itself. Bringing together a richer linguistic base with more subtle philosophical perspectives is crucial, the editors argue, to understanding the nature of causation in human cognition in general.

David Adger
Hagit Borer

Acknowledgements

This volume grew out of a conference called “Forces in Grammatical Structures” that was held in Paris in 2007. Initial versions of some of the chapters were presented there, while other chapters developed out of a conversation that began at that conference. We would like to thank the Centre National de la Recherche Scientifique, Université Paris 8, the laboratory Structures Formelles du Langage (UMR 7023), and the École Normale Supérieure for financial and logistical support for the conference, as well as the conference participants for their role in sparking the discussion that led to this book.

We are grateful to Clive Perdue for his work on both the conference and the volume; it was in fact his idea to make a volume out of the conference. Sadly, he passed away in the interim and was unable to see the end result of this project. We hope nonetheless that he would have been pleased.

The laboratory Structures Formelles du Langage and the Collaborative Research Center Incremental Specification in Context (SFB 732, Project B5) of the University of Stuttgart, financed by the Deutsche Forschung Gemeinschaft, provided instrumental support, financial and otherwise, for this volume.

We would also like to express our appreciation for the Fédération Typologie et Universaux Linguistiques, whose project “La causalité dans le langage et la cognition” provided financial support for the volume as well as the opportunity to run a forum for discussing causation. Thanks to Stéphane Robert and Véronique Brisset-Fontana at the Fédération, and Maya Hickmann and all the participants in the project.

We are most grateful to the authors whose contributions are represented here; we count ourselves fortunate to have had the opportunity to work with such excellent scholars in such a range of approaches to the topic of causation in language. We are likewise deeply appreciative of the reviewers whose insightful comments contributed greatly to the quality of the volume. Thanks too to Nigel Duffield and Heidi Harley for their help with initial preparations, and to the series editors Hagit Borer and David Adger.

The volume could not have been put together without the invaluable assistance of Sarah Barrett, Helen Bilton, Georgie Morand, Saralin Moser, Wendy Telfer, Madeleine Turcaud, and especially Alexandra Lehr.

Finally, many thanks to Julia Steer, Vicki Hart, and Kate Gilks, and to everyone else at OUP who has helped make this book a reality.

Notes on contributors

Bridget Copley is a senior researcher at the laboratory Structures Formelles du Langage, jointly affiliated with the Centre Nationale de la Recherche Scientifique and the Université Paris 8. Her research interests include causation, aspect, futures, and modality at the semantics-cognition and syntax-semantics interfaces. She received her Ph.D in 2002 from the Department of Linguistics and Philosophy of the Massachusetts Institute of Technology, and is the author of *The Semantics of the Future* (Routledge).

Nigel Duffield received his Ph.D in Linguistics from the University of Southern California in 1991; a revised version of his dissertation, on the syntax of Modern Irish, was published in 1995 (Kluwer). Since 1997, most of his syntactic research has focused on comparative aspects of Vietnamese grammar, with particular focus on the lexicon–syntax interface. He has held positions at various universities, including Heinrich Heine-University (Düsseldorf, Germany), McGill University (Canada), Radboud University (Nijmegen, The Netherlands), and the University of Sheffield (UK); currently, he is a Professor of Linguistics at Konan University (Okamoto, Japan).

Paul Egré is a CNRS research fellow and a member of the Institut Jean Nicod in Paris. His research is primarily in the areas of epistemology, philosophical logic, and the philosophy of language. His recent work focuses on the semantics and psychology of vagueness. One of his latest publications on this topic is the collective volume *Vagueness and Language Use* (coedited with N. Klinedinst, Palgrave Macmillan, 2011). Paul Egré is also the editor-in-chief of the *Review of Philosophy and Psychology*.

Raffaella Folli is currently a senior lecturer in the School of Communication at the University of Ulster, where she has been teaching and researching since 2005. Previously she held a research position at the University of Cambridge. She has two main areas of research: the lexicon–syntax interface (in particular verb alternations), on which she has published several articles in various journals; and the syntax of nominals cross-linguistically.

Heidi Harley is Professor of Linguistics at the University of Arizona. She is interested in the interfaces between lexical semantics and syntax, and between morphology and syntax, and has worked on these questions in English, Italian, Irish, Hiaki, Icelandic, and Japanese, largely within the Distributed Morphology framework. Her work has appeared in *Linguistic Inquiry*, *Language*, and the *Journal of Linguistics*, among others.

Tatjana Ilić recieved her Ph.D from the Department of Linguistics at the University of Hawaiʻi at Mānoa where she is currently affiliated as an adjunct assistant professor. Her research focuses on morphosyntactic and morpholexical operations, event structure, causation, agentivity, modality, language universals, and first language acquisition.

Max Kistler is Professor in the Department of Philosophy at Université Paris 1 Panthéon-Sorbonne and a member of the Institut d'Histoire et de Philosophie des Sciences et des Techniques, Paris, having previously held positions in Clermont-Ferrand, Paris-Ouest-Nanterre, and Grenoble. He is the author of *Causation and Laws of Nature* (Routledge, 2006; French edition 1999), co-editor (with B. Gnassounou) of *Dispositions and Causal Powers* (Ashgate, 2007), and editor of special issues of *Philosophie* (2006) on "Causation", *Synthese* (2006) on "New perspectives on reduction and emergence in physics, biology, and psychology", and *Philosophical Psychology* (2009) on "Cognition and neurophysiology: mechanism, reduction, and pluralism."

Anja Latrouite lectures in the Department of Language and Information at Heinrich-Heine-Universität Düsseldorf, and is a member of the SFB 991 research project B01, "Verb frames at the syntax–semantics interface." She wrote her Ph.D Thesis on voice and case in Tagalog. In addition to Austronesian languages, her research interests include phenomena related to argument linking and the syntax–semantics interface, verb semantics, and morphologically rich languages, as well as the evolution of morphology and syntactic structures.

Ekaterina Lyutikova is an Associate Professor at Lomonosov Moscow State University. Her specializations are generative syntax and the syntax–semantics interface. Current research areas include noun phase structure, syntax of relatives, argument structure, Turkic, North Caucasian, Iranian, and Uralic languages. She received her MA and Ph.D from Moscow State University. Recent publications include "Relative pronoun *kotoryj* in appositive relatives", in *Voprosy Jazykoznanija*; "Nominalization and the problem of indirect access: evidence from Ossetian", in *Linguistic Review*; and "Atelicity and anticausativization" (with Sergei Tatevosov), in *Argument Structure and Syntactic Relations: A Cross-linguistic Perspective* (John Benjamins).

Fabienne Martin is a research fellow with the project "Polysemy in a Conceptual System", which is part of the collaborative research centre (Sonderforschungsbereich) SFB 732 "Incremental Specification in Context" at the University of Stuttgart. She received her Ph.D in Linguistics in 2006 from the Université libre de Bruxelles. Her research interests include lexical semantics, aspect, as well as the syntax-semantics and semantics-pragmatics interfaces.

Gillian Ramchand pursues a research programme concerning the relationship between syntactic and semantic representations of natural language. She has worked in areas of tense, aspect, predication, and argument structure on languages as diverse as Bengali, Scottish Gaelic, and English. She is the author of two books, *Aspect and Predication* (OUP, 1997) and *Verb Meaning and the Lexicon* (CUP, 2008), where she argues for a syntactic implementation of an event-structural view of verbal meaning and participant relations. Ramchand is currently Professor of Linguistics at the University of Tromsø, Norway, and Senior Researcher at the Center for Advanced Study in Theoretical Linguistics (CASTL) there. She holds a Ph.D in Linguistics from Stanford University, and bachelor's degrees in Mathematics and in Philosophy from the Massachusetts Institute of Technology.

Florian Schäfer holds a research position with the project "Underspecification in Voice systems and the syntax–morphology interface", which is part of the collaborative research centre (Sonderforschungsbereich) SFB 732 "Incremental Specification in Context" at the University of Stuttgart. He studied General and Theoretical Linguistics at the University of Potsdam, and finished his dissertation on (anti-)causative alternation in 2007 at the University of Stuttgart. His main research interests are in the theories of syntax, morphology, and lexical semantics, and the interaction of these modules of grammar.

Sergei Tatevosov is an associate professor at Lomonosov Moscow State University. His specializations are morphology, formal semantics, and syntax–semantics interface. Current research areas include lexical and grammatical aspect, argument structure, Turkic, North Caucasian, Iranian, and Uralic languages. He received his MA and Ph.D from Moscow State University. Recent publications include "Telicity, measures, and enpoints", in *Proceedings of Sinn und Bedeutung* 16, "Detelicization and argument suppression: evidence from Godoberi", in *Linguistics* and "Atelicity and anticausativization" (with Ekaterina Lyutikova), in *Argument Structure and Syntactic Relations: A Cross-linguistic Perspective* (Benjamins).

Richmond H. Thomason is a logician with research interests in philosophy, linguistics, and computer science. His research interests include topics in philosophical logic, formal semantics and pragmatics, knowledge representation, and natural language processing. He received his Ph.D in Philosophy from Yale University in 1965.

Phillip Wolff received his Ph.D degree from Northwestern University, and is an Associate Professor of Psychology. He currently serves as director of the Cognition and Development program in Psychology, and is director of the joint major in Psychology and Linguistics. Wolff's research concerns the relationship between language and cognition, with particular interests in the representation of relational concepts, computational models of causal meaning and reasoning, and cross-linguistic approaches to the study of word meaning. He is the editor (with Barbara Malt) of the Oxford University Press volume *Words and the Mind*, and has served as associate editor for the journal *Cognitive Science*.

List of abbreviations

A	abilitative verb form
ABS	Absolutive case
ACC	Accusative case
ACH	Agent Control Hypothesis
ACT	Actuality aspect
ADESS	Adessive case
AG	Agentive
AgrO	Object Agreement
AgrP	Agreement Phrase
AP	Adjectival Phrase
ART	Article
Asp	Aspect
AspP	Aspect Phrase
AspP	Aspect Projection
AUX	Auxiliary
AV	Actor voice
BNC	British National Corpus
ç	Retroflex consonant
CAUS	Causative morpheme
CH	Aspirated consonant
CL	Classifier
COMPL	Completed
CONJ	Conjunctive subject
CONV	Converb
DAT	Dative case
DECL	Declarative
DEF	Definite determiner
Det	Determiner
DIR	Directive (full control)
DM	Distributed Morphology
DN	Deductive nomological
DP	Determinant phrase

DS	Different subject
e	Event
ECM	Exceptional Case Marking
ERG	Ergative case
ESP	Event-structural prominence
F	False
FEM	Feminine agreement
FI	*Faire* infinitif
FP	*Faire par*
FRUS	Frustrative morpheme
FUT	Future
GEN	Genitive case
HPSG	Head-driven Phrase Structure Grammar
IAC	Involuntary agent construction
IAsP	Inner Aspect
IC	INADVERTENT CAUSE
IMPF	Imperfective
INCR	Incremental
INF	Infinitive
Infl	Inflection
Init	Initiation
INSTR	Instrumental case
INTR	Intransitive
INUS	Insufficient-necessary/unnecessary-sufficient
IP	Inflection Phrase
IRR	Irrealis
ISP	Information-structural prominence
LC	Limited control
LCC	Lexical Causatives Constraint
LDG	Lexical Decomposition Grammar
LFG	Lexical Functional Grammar
LOC	Locative case
MASC	Masculine
N	Neutral verb form
NACT	Nonactive
NEG	Negation
NF	Non-finite

NMZ	Nominalization
NOM	Nominative case
NONFUT	Nonfuture tense
NONSQ	Non-sequence
NP	Nominal phrase
OAsp	Outer Aspect
PART	Partitive case
PASS	Passive morphology
PAST/PST	Past tense
PERF	Perfective
PL	Plural
POSS	Possessive
PP	Prepositional phrase
PRES	Present tense
PRN	Pronoun
Proc	Process
PROG	Progressive
PRS	Present
PTCL	Particle
PTCP	Participle
R	Relation
REFL	Reflexive marker
Res	Result
RL	Realis
RP	Referential prominence
RRG	Role and Reference Grammar
S/SG	Singular
Spec	Specifier
SU	Indicative subject clitic
SV	Subject Verb
T	True
θ_{EXT}	External argument
TH	Theme
TOP	Topic
TR	Transitive
UMC	Unique Mapping Constraint

UV	Undergoer voice
Ṽ	Nasalized vowel
V	Verb
VP	Verb phrase
VRB	Verbalizer

1

Introduction

BRIDGET COPLEY AND FABIENNE MARTIN

The linguistic expression of causation has been a topic of interest to philosophers since antiquity. With the burgeoning of the science of linguistics in more recent times, linguists have also begun to be interested in causation, recognizing the important role it plays in diverse linguistic phenomena. Causation is most obviously referred to directly through lexical causatives such as *cause*, but it is relevant to the meanings of many other elements in language. For example, a causal relation has been proposed as a part of the meaning of accomplishments such as *build a house* or *bake a cake* (Pustejovsky 1995; Higginbotham 2009; Ramchand 2008). The notion of intentional causation in the domain of agency is crucial to the understanding of verbal syntax and semantics (DeLancey 1984; Levin and Rappaport Hovav 1995; Reinhart 2002; Folli and Harley 2005; Alexiadou et al. 2006). As Talmy's (e.g. 1988; 2000) force-dynamic view of meanings suggests, the meanings of modals such as *can* and *must*, and the meanings of verbs such as *help* and *prevent*—and indeed all verbs, a view furthered by Croft (e.g. 1991) and Gärdenfors (e.g. 2000)—are plausibly related to our understanding of causation. And in less-studied languages, the issues of unintentional causation, unachieved goals, and other interesting kinds of causation have come to the forefront (e.g. Zepeda 1987; Travis 1991; 2000; Davis and Demirdache 2000; Pylkkänen 2002; Jacobs 2011; Fauconnier 2012).

Discussion of such causally related phenomena has been lively within cognitive linguistic approaches and typological studies (e.g. Song 1996; Shibatani 2002), as well as in more structurally oriented work in the generative tradition. It is, however, becoming increasingly clear that if we are to further our understanding of how causal meanings are represented in language, it will be important to take advantage of all the resources at our disposal.

For example, we see already in this volume that causal phenomena in less-studied languages are important to theories of causation in language. The idea that cross-linguistic data is important will be uncontroversial to linguists, but it is worth underlining, since such typological and cross-linguistic research is methodologically very different from the philosophical tradition of relying on the philosopher's own intuitions. Any theory of language must come to terms with the wide, yet relatively constrained variation in form and meaning seen among the world's languages, some of which may contradict received philosophical consensus (see Copley and Wolff, Ch. 2, this volume). To the extent that philosophers and cognitive scientists make claims about language, they must also come to terms with cross-linguistic variation, especially the variation seen in less-studied languages.[1] Among the languages discussed in this volume are several less-familiar languages such as Tohono O'odham, Finnish, Yup'ik Eskimo, Hindi/Urdu, Vietnamese, Karachay-Balkar, and Tagalog.

In addition to taking advantage of the wealth of causal expressions and phenomena in the languages of the world, the nature of semantics itself demands that in order to further our understanding of causal meanings, a wide range of causal theories must be considered. As Chomsky (1995) famously reminds us, semantics has two interfaces: the conceptual system and the computational system (i.e. syntax). If theories of the semantics of causation are to continue to advance, research on causal concepts—both philosophical and psychological—must be combined with syntactic theory of the structures they are embedded in. In other words, theories of causal meaning, or indeed any meaning, must be both conceptually plausible and syntactically honest. It is interesting to note that in this position there is no basic conflict with cognitive linguistics, which has always taken seriously the interface between cognition and language, and which can be seen as a good starting point for the bringing together of conceptual and syntactic theory. In a way, the promise of the cognitive linguistic approach is fulfilled in an approach that sees semantics as the bridge between purely conceptual and purely structural evidence.

With this background in mind, the question arises as to how far along we are in the quest to understand the semantics of causation via its two interfaces: the cognition–semantics interface and the syntax–semantics interface.

As far as the cognition–semantics interface is concerned, the integration between theories of the cognition of causation and theories of language is nearly nonexistent. While there has never been universal consensus as to the nature of causation, we will see in this volume—especially in Copley and Wolff's chapter and the other chapters of the first section—that this debate should be of interest to linguists. Moreover, it is a good time to be interested in theories of causation, as the traditional positions in the debate have recently been pursued intensively to a much more advanced level than has ever been seen before. A burst of philosophical interest in causation over the last ten years has led to highly sophisticated accounts of causal phenomena, with the hypothesis space dramatically reduced. Likewise, a trend toward formalization makes these theories at once more falsifiable and, potentially, more accessible to linguists.

[1] We can only speculate on how philosophical theories of causation might have developed differently through the ages if their proponents had been speakers of non-Indo-European languages.

In the fields of psychology and computer science, too, there have been a number of recent advances in the analysis of causation. In particular, cognitive scientists have shown how causal relationships can be used to make forward and backward inferences about the probability of events, including events embedded in complex networks of causation (e.g. Pearl 2000). Further, there has been real progress in investigating the cognitive plausibility of proposed models of causation (e.g. Sloman 2005; Gopnik and Schulz 2007; Marcus and Davis 2013).

All of this is to indicate that access to the conceptual interface with semantics in the discussion of linguistic expression of causation is not only desirable but possible. An additional reason why it is possible is that generative linguistics has matured as a field. Over the last thirty years, as generative work at the syntax–semantics interface has become more and more refined in mapping meanings to syntactic components, linguists have been asking ever more detailed questions about the meanings of these components. At the same time, the complexity that was characteristic of early transformational grammar has been winnowed down, in the Minimalist Program (Chomsky 1995), to a single compositional operation MERGE. Likewise, there is a concerted drive to simplify the inventory of semantic combinatorial operations (see e.g. Heim and Kratzer 1998). Such simplification makes linguistic theory more cognitively realistic, and thus more open to interaction with theorists in related disciplines.

This brings us to the second interface, that of semantics with syntax. The last twenty years have seen a veritable explosion of work at this interface, a good deal of it concerning causation. On the generative approach to language, theories of meaning are understood to be constrained by syntactic considerations, so there arises in the generative approach a possibility of better understanding how language represents causation by coming to see how causal meanings are syntactically structured.

Despite the fruitfulness of this approach, there is still a need to bring structural approaches to causation more closely together, and this constitutes another reason why we have been interested in bringing the two sections of this volume together into one conversation. It seems to us that linguists interested in lexical semantics study causation in lexical semantics; those interested in causation and agency study these topics in argument and event structure; and causation as it pertains to modality, aspect, and other phenomena has been addressed only in passing.

At each interface, then, particular points need addressing with respect to causal meaning: the cognition–semantics interface should be informed by the philosophical and psychological perspectives on causation, and the syntax–semantics interface should see greater integration across linguistic phenomena related to causation. Finally, both interfaces must ultimately relate to each other through the semantics, in a theory of meanings of causal (cross-)linguistic phenomena. The chapters in this volume are offered in this context.

1.1 From causal concepts to causal meanings

Part I of the volume deals with causation at the cognition–semantics interface, i.e. the mapping from conceptual representations of causation to the representations of causal meaning in language.

In Chapter 2 ("Theories of causation should inform linguistic theory and vice versa"), **Bridget Copley** and **Phillip Wolff** offer a basic introduction to the different approaches philosophers take to causation. These approaches may be divided into two categories: dependency theories, in which a cause C causes an effect E just in case E depends on C in some way (familiar to linguists through David Dowty's 1979 adaptation of David Lewis's 1973 theory of causation), and production theories, in which C causes E just in case a certain configuration of influences holds of C and E, or some conserved quantity is transmitted from C to E. Copley and Wolff argue that a familiarity with these theories would be fruitful for linguists working on causation in language, and give examples (defeasible causation, volitionality, and causal chain mappings) where the choice of causal theory has ramifications for the linguistic theory; they also contend that linguistic theory has the potential to inform philosophers and cognitive scientists working on causation as well.

Chapter 3 ("Formal semantics for causal constructions"), continues the effort to bring together philosophers and linguists. In this chapter, **Richmond Thomason** addresses Dowty's extension of Montague's Intensional Logic to the problem of causation. He identifies some difficulties that arise in Dowty's approach, and suggests an alternative event-based theory that, while it does not provide a global interpretation of causality, seems to work well with a wide range of the causal constructions that are important in word formation. The notions of telicity, agency, and direct vs. indirect causation are all addressed in Thomason's proposal. He further relates these ideas to normality and the formalization of commonsense defeasible reasoning.

Max Kistler, in Chapter 4 ("Analyzing causation in light of intuitions, causal statements, and science"), suggests that the existence of two equally plausible but incompatible approaches to causation has its source in the conflict between two types of intuition. Some causal judgments are justified by the intuition of nomic dependency, i.e. dependency of one state of affairs on another by virtue of laws of nature. Other causal judgments are made on the basis of a material influence or transmission between events. These two types of intuition lie behind the tension between an explanatory concept and a mechanistic conception of causation. Kistler first argues that causal statements relating facts express the explanatory aspect of causation, and causal statements relating events express the mechanistic aspect. Relying in part on nominalization data, Kistler goes on to propose a framework that reconciles the two aspects and shows the logical relations between statements of the two sorts. Finally, he analyzes certain types of causal statements that do not seem to fit in the proposed scheme: statements expressing interruption, triggering, and omission, where counterfactual information about what would have happened "normally" must be taken into account, where normality may be understood in terms of statistical average, biological fitness, or morality.

Chapter 5 ("Causal pluralism and force dynamics") represents another perspective on the distinction between the two different kinds of theories of causation, rejecting the idea that both kinds of theory are necessary. **Phillip Wolff** argues that dependency models represent causation in terms of kinematics, i.e. with respect to the observable properties of events. In contrast, his dynamics model (a kind of production theory which is based on Talmy's 1988 theory of force dynamics), specifies causation in terms of dynamics: the invisible quantities that produce kinematic patterns. In the dynamics model, causation is characterized as a pattern of forces and a position vector. This model is supported by studies in which participants watched 3D animations generated from a physics simulator. In these experiments, the very same forces used to generate physical scenes were used as inputs into a computer model to predict how those scenes would be described. In a second line of experiments, the model is extended to sequences of events in which configurations of forces are linked together by their resultant vectors. The model was able to predict when a causal chain could be described in more than one way, and to what degree. Thus, unlike any other model to date, the dynamics model offers an explanation of the relationship between deterministic and probabilistic causation, as well as of the semantics of several complex predicates.

One obstacle to applying production theories of causation to formal semantic theory is that forces are typically not represented in formal semantics. **Bridget Copley** and **Heidi Harley** point out in Chapter 6 ("Eliminating causative entailments with the force-theoretic framework: The case of the Tohono O'odham frustrative *cem*") that in many cases in natural language, causation must be treated as "defeasible"—i.e. one event is asserted or presupposed to normally cause a second event, but there is no entailment that the second event actually occurs. To account for such cases, they propose that the arguments discovered by Davidson refer to forces instead of to events. A force, conceptually, is energy input into a situation. Formally, Copley and Harley treat forces as functions from an initial situation to the situation that results *ceteris paribus* (all else being equal). This allows for the possibility that all else may not be equal, leading to the lack of a causative entailment. Copley and Harley illustrate the framework with an analysis of the frustrative morpheme *cem* in Tohono O'odham, a Uto-Aztecan language spoken in southern Arizona and northern Mexico. The resulting analysis sheds light on statives and the nature of plans, as well as on prospective, imperfective, and perfective aspect.

Chapter 7 ("Modality and causation: Two sides of the same coin") also relates the production perspective on causation to language, this time from a cognitive/typological perspective. On the basis of cross-linguistic data from languages such as Yup'ik Eskimo, Italian, Serbian, and Finnish, **Tatjana Ilić** argues that modal and causative meanings are fundamentally related through the notion of control over the event, which itself is best understood through the lens of a production theory of causation. Causative meanings, she proposes, arise in causal chains with agentive verb bases whose subject is devoid of control over the event. Typically, such a chain also involves an initiator—not the subject—who initiates the event and controls its outcome. In this type of chain, modal presuppositions of obligation on the causee can arise. In contrast, when a causal chain with an agentive verb base and the agent argument does not involve an initiator who can initiate the event and control its outcome, the chain fails to obtain causative interpretation. In this type of chain, modality surfaces as the asserted meaning, replacing the meaning of causation.

Paul Egré, in Chapter 8 ("Intentional action and the semantics of gradable expressions (On the Knobe Effect)"), examines an hypothesis put forward by Pettit and Knobe (2009) to account for the Knobe Effect, i.e., the fact that speakers are more or less likely to judge actions as intentional depending on certain circumstances of the action. According to Pettit and Knobe, one should look at the semantics of the adjective *intentional* on a par with that of other gradable adjectives such as *warm*, *rich*, or *expensive*. What Pettit and Knobe's analogy suggests is that the Knobe Effect might be an instance of a much broader phenomenon which concerns the context-dependence of normative standards relevant for the application of gradable expressions. Egré adduces further evidence in favor of this view, and goes on to examine the predictions one obtains when assuming that *intentional* involves a two-dimensional scale, delimiting how much an action or outcome is desired on the one hand and how much it can be foreseen as a consequence of one's actions on the other.

1.2 From causal meanings to causal structures

The chapters in Part II investigate causation at the syntax–semantics interface—the mapping between causal meanings and the syntax by which these meanings are structured.

Fabienne Martin and **Florian Schäfer** present Chapter 9 ("Causation at the syntax–semantics interface") as an overview of the themes that arise in the study of causation at the syntax–semantics interface. They present and discuss recent proposals about the argument structure and the event decomposition of (anti-)causative verbs, and illustrate the deep interconnection between these two layers through several generalizations that have been put forth in the recent literature: (i) causer (i.e. inanimate) subjects require a resultative event structure; (ii) non-culminating readings of accomplishments require the predicate's external argument to be associated with agentive properties; and (iii) the difference between agent vs. causer subjects affect transitivity. Most of the phenomena addressed illustrate the crucial role played by the thematic properties of external arguments in the syntax and semantics of causative verbs, as well as the importance of a more fine-grained typology of agents (intentional or not, endowed with control over the action or not) for the syntax. Additionally, Martin and Schäfer address some of the differences between mono-clausal vs. bi-clausal causatives, focusing on the distinction between indirect and direct causation. Again, the thematic properties of external arguments are shown to play a crucial role in this distinction too (and more particularly in the possibility to use a lexical causative to express indirect causation).

In Chapter 10 ("Causal chains and instrumental case in Hindi/Urdu"), **Gillian Ramchand** revisits the licensing and interpretation of instrumental case-marked nominals in Hindi/Urdu causative constructions. In these constructions, the instrumental *se*-marked adjunct is licensed with an "intermediate agent/causee" interpretation in the indirect morphological causative using the suffix *-vaa* (Masica 1991; Saksena 1982b; Kachru 1980; Hook 1979), inviting comparisons with the demoted agent analysis of English *by*-phrases (Jaeggli 1986; Grimshaw 1990; Baker et al. 1989; Embick 2004). Ramchand argues against the hypothesis that the *se*-marked phrase corresponds to a demoted agent. Rather, she argues, a more unified analysis of *se*-phrases can be achieved through an event-structural analysis, in line with the standard interpretation of other adverbials in the syntax (cf. Ernst 2002). Since the "intermediate agent" interpretation is only possible with indirect causatives in Hindi/Urdu, the event-structural analysis proposed here also has implications for the direct vs. indirect causation distinction in the syntax.

Ekaterina Lyutikova and **Sergei Tatevosov** argue in Chapter 11 ("Causativization and event structure") that Pylkkänen's (2002) comprehensive theory of the causative at the syntax–semantics interface in a number of languages faces complications. The authors point out that there are languages, one of which is Karachay-Balkar (Altaic, Turkic), for which the theory does not always make correct predictions. The main goal of their chapter is thus to develop an alternative that incorporates new data and accounts for syntactic and semantic characteristics of causatives in languages like Karachay-Balkar. First, the authors challenge Pylkkänen's suggestion that the causative falls under exactly one of the three structural types, Root-selecting, Verb-selecting, and Phase-selecting. Secondly, an account is presented for the semantic distinction between direct and indirect causatives, problematic for Pylkkänen. A novel architecture of the verbal domain is proposed whereby relations between subevents in a syntactically represented event structure are introduced independently from subevent descriptions.

In Chapter 12 ("Inadvertent cause and the unergative/unaccusative split in Vietnamese and English"), **Nigel Duffield** draws together several strands of evidence in support of the claim that two kinds of cause relation are independently represented in phrase structure. The first of these is the familiar intentional/volitional cause associated with the thematic relation AGENT, typically represented in the current generative literature as the argument licensed by "little *v*": in recent years, it has once again become commonplace to assume that this intentional CAUSE is abstractly represented in phrase structure, either as a primitive predicate or as a relational notion; see Hale and Keyser 1993; Baker 1997; also Pustejovsky 1991; Tenny and Pustejovsky 2000b. Duffield's chapter, however, focuses on the structural representation of the second type of cause: a less studied relation Inadvertent Cause (IC), which—in contrast to its more robust cousin—has escaped detailed scrutiny until quite recently. The analysis presented here develops a proposal originally articulated by Travis (1991; 2000; 2010), which associates the IC thematic relation with the specifier position of a VP-internal functional category, namely, Inner Aspect (IAspP). Travis' proposal is originally motivated by facts from a completely different range of (Western Malayo-Polynesian) languages: to the extent that it extends naturally to the phenomena discussed here, the present work provides confirmation of the profitability of a syntactic approach to inadvertent cause.

In Chapter 13 ("Causatives and inchoatives in the lexicon and in the syntax: evidence from Italian"), **Raffaella Folli** revisits an ongoing debate on the causative/inchoative alternation. In this debate there are two opposing approaches: one assuming that the alternation is due to causativization in the syntax of an underlying basic unaccusative structure (e.g. Embick 1997; Folli 2001; Harley 1995; Ramchand 2008), and one arguing that a lexical operation of decausativization or reduction of the external causer argument is responsible for the inchoative form (e.g. Levin and Rappaport-Hovav 1995; Chierchia 2004; Reinhart and Siloni 2005; Koontz-Garboden 2009). Folli presents data from Italian to argue that this language supports a more flexible approach to the derivation of this kind of alternation. In particular, she argues that Italian distinguishes three classes of verb that participate in this alternation, and that in fact for two classes of change-of-state verbs, the first type of syntactic operation described above is at work in the formation of causative forms, while for another class the alternation is lexical.

Finally, **Anja Latrouite** investigates Tagalog argument realization patterns in Chapter 14 ("Event-structural prominence and forces in verb meaning shifts"), with a special focus on the construal of events based on subject choice. Tagalog, a Philippine Austronesian language classified typologically by Drossard (1984) as an 'active' language, makes explicit aspects of causal structure not visible in languages of the European type. Tagalog is known for its complex verbal affixation, as well as the fact that almost every argument in a sentence can be the subject, and may be marked on the verb by a voice affix signalling its thematic role. However, it has been observed that there are restrictions on subject selection, and that voice choice may lead to shifts in the interpretation of verbs. This chapter explores these restrictions and related meaning shifts, and argues that voice selection is based on a number of prominence considerations on different levels. On the level of event structure, prominence is shown to be tightly linked to disparate elements associated with the causal construal of the event. Thus, the nature of events, the properties of the participants involved, and the relation between them is shown to play a central role in the overall grammatical system.

Part I

From Causal Concepts to Causal Meanings

2

Theories of causation should inform linguistic theory and vice versa

BRIDGET COPLEY AND PHILLIP WOLFF

2.1 Introduction

Linguistics has long recognized that causation plays an important role in meaning. Over the last few decades of the generative linguistic project, it has become clear that much of phrase structure is arranged around causal relationships between events (or event-like entities such as situations). Reference to causation in this tradition has most often taken the form of a relation CAUSE, with little further elucidation, in effect treating CAUSE as a primitive. This treatment of causation as a primitive relation has proved adequate to the task of developing grammatical structures that make reference to causation. But arguably, this hands-off approach to the meaning of causation has obscured potentially relevant details, impeding linguists' ability to consider hypotheses that might yield a more comprehensive analysis of the roles played by concepts of causation in language. Unpacking the notion of causation should, on this view, afford a deeper understanding of a range of linguistic phenomena, as well as their underpinnings in conceptual structure.*

In this chapter, we show how attention to the variety of existing theories of causation could advance the understanding of certain linguistic phenomena. In the first section, we review the two major categories of theories of causation, including some of the principal challenges that have been raised for and against each category. We identify in the second section a range of linguistic phenomena that we feel would benefit from a deeper investigation into causation—defeasibility, agentivity and related concepts, and causal chains—and also speculate on how theories of causation might inform our understanding of these phenomena. Since the linguistic theories make testable claims about cognition, they give rise to potential connections between syntactic structure and cognition. In the concluding section, we express our hope that further investigations along these lines

* Thanks to Kevin Kretsch and Jason Shepard for helpful discussion.

may pave the way for a theory of meaning grounded in both syntactic and cognitive realities, in a way that has not previously been possible.

2.2 Theories of causation

Given the importance of causation in linguistic theory, the question naturally arises as to whether some of the varied insights about causation from philosophy and cognitive psychology might have consequences for our understanding of linguistic theory. Up until now, they have not. The single theory of causation most often referred to in linguistic articles—by an extremely large margin—is Lewis's (1973) counterfactual theory of causation, which is discussed and adapted in Dowty's influential (1979) book. But even when Lewis or Dowty are cited, the causal relation is usually[1] treated by linguists essentially as a primitive. As a consequence, even Lewis's theory has not had a particularly meaningful impact on our understanding of the role of causal concepts in forming causal meanings.[2] Certainly, when causation has been treated as a causal primitive, it has just been a placeholder, a way of not having to deal with what causation is. Historically, this move was defensible since it was not clear that linguistic phenomena really depended on how causation was defined, or whether the grammar had access to anything more fine-grained than a primitive relation CAUSE. Arguably, it was even provisionally necessary to treat causation as an unanalyzed primitive at the outset of the development of the syntax–semantics interface, to avoid unnecessary complication.

As the generative enterprise has progressed, however, the need to address the lacunae still present in linguistic phenomena related to causation has become more and more pressing, both in familiar and in novel data. A number of linguistic phenomena, some of which we will present in this chapter, are not well addressed by appeal to a primitive CAUSE. It has therefore become increasingly apparent that

[1] Notable exceptions—i.e. authors who further investigate the Lewis–Dowty approach—include: Bittner (1998) (type lifting for cases where causal meaning is morphologically unmarked); Eckardt (2000) (focus sensitivity of the verb *cause*); Kratzer (2005) (a causal head in resultatives); Neeleman and van de Koot (2012) (questioning whether there is a causal event associated with causative predicates; see also Van de Velde 2001 for a similar point); Truswell (2011) (constraints that causal structure puts on extraction). A related theory, that of causal modeling (see section 2.2.1.4) is starting to be of interest to people working on modals and counterfactuals; e.g. Dehghani et al. (2012). Production theories of causation that rely on forces, or transmission of energy are very similar to a parallel development in cognitive linguistics that had its start with Talmy (1985a; 1985b; 1988). A few lines of inquiry in formal linguistics have explicit, or implicit links to production theories, most notably those of van Lambalgen and Hamm (e.g. 2003; 2005) and Zwarts (e.g. 2010).

[2] Lewis's theory has had an enormously meaningful and fruitful impact on semantics in the realm of conditionals (counterfactual and otherwise) and modals, stemming from initial work by Stalnaker (1968 and much later work), and Kratzer (1977; 1979 and much later work), as well as Dowty's work on the progressive (1977; 1979). The clear predictions and expressive power of the possible worlds approach have deservedly made it a jewel in the crown of modern semantic theory. However, this body of research has not generally been explicitly linked to the issue of causation. As we will discuss in section 2.2.2, causation and at least one kind of modal notion (that of volitionality) are related to each other, whether or not one agrees with Lewis on the best way to represent them; see also Ilić, Ch. 7, this volume, for discussion of linguistic data bearing on the relationship between causality and modality.

the story one tells about causal meanings will have to depend on one's theory of causal concepts. If, for instance, we were to take the details of Lewis' counterfactual theory of causation into account, it could have some interesting consequences.

Lewis's theory, though, is not the only theory in town. Other modern theorists have suggested that causal relations might be based on statistical dependencies (Suppes 1970; Eells 1991; Cheng and Novick 1991), manipulation (Pearl 2000; Woodward 2003), necessity and sufficiency (Mill 1973; Mackie 1974; Taylor 1966), transfer of conserved quantities (Dowe 2000), force relations (Fales 1990; White 2006), energy flow (Fair 1979), causal powers (Mumford and Anjum 2011a), and property transference (Kistler 2006a). While this list accurately reflects the considerable variation among philosophers as to the nature of causation, discussions of causation often categorize theories of causation according to several dimensions, such as whether the relata are single or generic, individual or population level; whether the causal relation is physical or mental; whether the causal relation is objective or subjective; or whether it is actual or potential (see Williamson 2009, e.g.). Many of these distinctions are not particularly relevant to current linguistic understanding of the causal relation as an element that occurs in a wide range of different environments. For example, linguistic consensus treats genericity as a separate operator from the causal relation CAUSE, so any viable proposal for the latter must be consistent with both generic and individual causation. In the following categorization of causation we emphasize two broad categories: *dependency* theories, in which A causes B if and only if B depends on A in some sense, and *production* theories (also commonly referred to as *process* theories), in which A causes B if and only if a certain physical transmission or configuration of influences holds among the participants in A and B.

2.2.1 Dependency theories

One major category of theories holds that causation is understood as a *dependency*. There are three main classes of dependency theory.

2.2.1.1 Logical dependency There is intuitive appeal in defining causation in terms of necessary and/or sufficient conditions. However, an analysis of such accounts raises a range of problems that are generally considered insurmountable (see Scriven 1971; also Hulswit 2002; Sosa and Tooley 1993). Consider, for example, a definition that identifies the concept of causation as a condition that, in the circumstances, is necessary. With such a definition, we might agree with Hume's comment that one event causes another "where, if the first object had not been, the second never had existed" (Hume 2007[1748])—i.e. a cause is a factor without which the effect would not have occurred.[3] The simplest version of such an account is contradicted by cases of late

[3] "Necessary" here is not necessarily to be thought of in the later modal logic sense of quantification over possible worlds; see e.g. Hume's "necessary connection" between cause and effect wherein "the determination of the mind, to pass from the idea of an object to that of its usual attendant" (Kistler 2006a).

pre-emption, i.e. cases where a potential alternative cause is interrupted by the occurrence of the effect.[4] For an example of late pre-emption, consider a scenario developed by Hall (2004).

There is a bottle on the wall. Billy and Suzy are standing close by with stones and each one throws a stone at the bottle. Their throws are perfectly on target. Suzy happens to throw first and hers reaches the bottle before Billy's. The bottle breaks. In this scenario, the effect of a particular candidate cause, Billy's throw, is "pre-empted" by another cause, Suzy's throw. As empirically verified by Walsh and Sloman (2005), Suzy's throw is understood to be the cause of the bottle's breaking, but Suzy's throw was not a *necessary* condition for the effect: if Suzy had missed, the bottle still would have broken because of Billy's throw.

An alternative account of causation in terms of logical dependency would be the proposal that causation is a sufficient condition for an effect. Under this view, a factor is the cause of an effect if the presence of that factor guarantees the occurrence of an effect. Of course, one problem with this view is that it is rare to find a case where single condition is sufficient in and of itself. An event is rarely, if ever brought about by a single factor; as Mill (1973[1872]) notes, every causal situation involves a set of conditions, which are sufficient for an effect when combined. Another problem for a sufficiency view is the case of late pre-emption described above. As noted, we would not say that Billy caused the breaking of the bottle. This is surprising from a sufficiency view, since Billy's throw is a sufficient condition for the breaking of the bottle.

Yet another possibility would be to define a cause as a necessary *and* sufficient condition (Taylor 1966). Such a definition fails because it entails that the cause would be a necessary condition and, as already discussed, there can be causes that are not necessary. A related view of causation is Mackie's (1965) INUS condition, that says a cause is an *insufficient* but *necessary* part of a condition which is itself *unnecessary* but *sufficient* for the result. The INUS condition, ultimately, defines causation in terms of sufficiency, but as discussed above, a factor (or set of factors) can be sufficient and yet not be a cause. A modern instantiation of an account of causation based on logical necessity and sufficiency can be found in Goldvarg and Johnson-Laird's (2001) model theory.

2.2.1.2 Counterfactual dependency Another type of dependency theory is based on the idea of counterfactual dependency: the counterfactual proposition that E would not have occurred without C. As we have seen, counterfactual dependency can be thought of as a paraphrase of the proposition that C is necessary for E. The modus operandi behind counterfactual theories of causation is thus to link two groups of

[4] Late pre-emption occurs when there are two potential causes but the occurrence of the effect prevents one of the causes from causing the effect. Early pre-emption (to be discussed in section 2.2.1.2) occurs when the initiation of one cause prevents the other potential cause from happening at all. See Menzies (2008) and Paul (2009) for more details.

intuitions: intuitions about whether certain counterfactual propositions are true and intuitions about whether certain events cause other events.

The simplest way to link these intuitions would be to identify causation with counterfactual dependency: i.e. to say that C is a cause of E if and only if E would not have occurred if C had not occurred. This looks as though we are equating causation with logical necessity, because it asserts that C must be present in order for E to occur. As we have seen, a definition of causation in terms of logical necessity erroneously predicts that C is not a cause of E if E could have been caused by something other than C. David Lewis, in the original version of his influential counterfactual theory of causation (1973 et seq.), proposed to avoid this problem by weakening the biconditional ("if and only if") to a mere conditional: counterfactual dependency entails causation, but causation does not entail counterfactual dependency. According to Lewis (1973), the reason that causation does not entail counterfactual dependency is because causal relations can sometimes emerge from transitive reasoning, but counterfactual relations, arguably, are not transitive (see Stalnaker 1968), and so causal relations may sometimes exist in the absence of a counterfactual dependency.

An example of such a scenario occurs in cases of so-called *early* pre-emption. Imagine, for example, a slightly different version of the Billy and Suzy scenario that was discussed above (which demonstrated *late* pre-emption). In this new scenario, Suzy throws a rock at a bottle (breaking it) and Billy acts as a backup thrower just in case Suzy fails to throw her rock. Here Suzy is the cause of the bottle's breaking, but just as in the case of late pre-emption there does not exist a counterfactual dependency between Suzy and the bottle's breaking; if Suzy had not thrown, the bottle would have still been broken because Billy would have thrown his rock.

To insulate his theory against such scenarios, Lewis (1973) proposes that C causes E if and only if *stepwise* counterfactual dependency holds between C and E, i.e. only if there are counterfactual dependencies holding between adjacent events in the chain, but not necessarily non-adjacent events in the chain. In the early pre-emption scenario, Lewis (1973) would argue that while the rock's breaking does not depend counterfactually on Suzy, there is a counterfactual dependency between Suzy and the intermediate event of the rock flying through the air, and a counterfactual dependency between the rock flying through the air and the bottle's breaking, and this chain of counterfactual dependencies licenses a judgment that Suzy's throw caused the bottle to break.[5] Lewis' approach to the problem raised by early pre-emption ultimately led to a definition of causation in terms of causal chains: specifically, C is a cause of E if and only if there exists a causal chain leading from C to E. Importantly, however, the links in the causal chain are defined in terms of counterfactual dependencies.

[5] Contra Lewis (1973), it is not entirely clear that there exists a counterfactual dependency between Suzy's throw and a rock flying through the air. Had Suzy not thrown her rock, there still would have been a rock flying through the air due to Billy.

There are a number of problems with Lewis's initial proposal, some of which continue to complicate counterfactual theories today. One kind of problem occurs in the case of *late* pre-emption. In both early and late pre-emption, a counterfactual dependency fails to hold between C and E, suggesting that counterfactual dependency is not *necessary* for causation. Lewis (1973) was able to address the lack-of-necessity problem in cases of early pre-emption by defining causation in terms of *stepwise* counterfactual dependency; but this fix only works for early pre-emption, not for late pre-emption, so the lack-of-necessity problem remains in the case of late pre-emption. Two other problems can be illustrated with a single type of scenario (see Hall 2000). Consider a case where an assassin places a bomb under your desk, causing you to find it, which causes you to remove it, which causes your continued survival. Without the assassin putting the bomb under your desk, you would not have removed it and thereby ensured your survival. Cases such as this demonstrate that counterfactual dependencies are not *sufficient* for causation. In this example, there exists a counterfactual dependency between the assassin and survival, but we would not want to say that the assassin caused your continued survival. Such cases also raise a problem for Lewis' (1973) definition of causation in terms of causal chains. As already noted, this definition was motivated by the assumption that causation is transitive; but, as shown in this example, there may be cases where transitivity in causation fails (see also McDermott 1995; Ehring 1987).

Lewis's 2000 theory attempts to address several of the problems facing his 1973 theory. In Lewis's new theory, counterfactual dependency exists when alterations in the cause lead to alterations in the effect. So, for example, if Suzy's throw is slightly altered—she throws the rock a bit faster, or sooner, or uses a lighter rock—the resulting breaking of the bottle will also be slightly altered. Lewis's new theory is able to explain why Suzy's throw, and not Billy's throw, is considered to be the cause: alterations to Suzy's throw result in changes in the effect, while alterations to Bill's throw do not. However, there is reason to believe that Lewis's new theory still does not escape the challenge raised by late pre-emption. As noted by Menzies (2008), in the case of Billy and Suzy, there is a degree to which alterations in Billy's throw could result in alterations of the final effect—if, for example, Billy had thrown his rock earlier than Suzy's. In order for Lewis's theory to work, only certain kinds of alteration may be considered. To foreshadow a point we will later make in the discussion of production theories (section 2.2.2), it may be that Lewis's theory can be made viable if the alterations are confined to those that are relevant to the creation of forces.

2.2.1.3 Probabilistic dependency According to Hume (2007[1748]), if it is true that an event C causes an event E, it is true that events similar to C are invariably followed by events similar to E. This view is referred to as the "regularity theory" of causation. A well-known difficulty with the regularity theory is the simple observation that causes are *not* invariably followed by their effects. The observation has motivated accounts of causation that ground the notion of causation in terms of probabilistic dependency.[6]

The simplest type of probabilistic dependency is one that relates causation to probability raising (Reichenbach 1956; Suppes 1970; Eells 1991). A variable C raises the probability of a variable E if the probability of E given C is greater than the probability of E in the absence of C (formally, $P(E \mid C) > P(E \mid \neg C)$). Thus on this theory, if smoking causes cancer, the probability of cancer given smoking is greater than the probability of getting cancer in the absence of smoking. An alternative way of describing the relationship between the conditional probabilities $P(E \mid C)$ and $P(E \mid \neg C)$ is to say that C is a cause of E when C makes a difference in the probability of E. Indeed, whenever $P(E \mid C) > P(E \mid \neg C)$ holds, E and C will be positively correlated and whenever E and C are positively correlated, $P(E \mid C) > P(E \mid \neg C)$. A relatively recent instantiation of probability raising is instantiated in Cheng and Novick's (1992) probabilistic contrast model.

While probabilistic approaches to causation address important limitations not addressed by other dependency accounts, they do not escape some other problems. Probability raising on its own seems to be not *sufficient* for causation: that is, C might raise the probability of E without C's being a cause of E. The reason it is not sufficient is because the presence of one event might (appear to) make a difference in the probability of another, but that appearance might in fact be due to a shared common cause, rather than from one causing the other (Hitchcock 2010). So, for instance, seeing a spoon raises the probability of seeing a fork; not because spoons cause forks, but rather because there is some overlap between the causes of seeing a spoon and the causes of seeing a fork. Reichenbach suggested that such cases could be flagged in the following manner: if two variables are probabilistically dependent and if one does not cause the other, they have a common cause that, if taken into account, renders the two variables probabilistically independent. Williamson (2009), however, points out that Reichenbach's characterization excludes cases of probabilistic dependency where C and E are related logically, mathematically, through semantic entailment, or accidentally.

Even with Reichenbach's common-cause cases excluded, however, sufficiency is still a problem. Returning to Suzy and Billy's case of late pre-emption, we can also see that probabilistic dependency is not sufficient for causation (C raises the possibility of E but C is not a cause for E). We can imagine that Billy's throw hits the bottle with a certain probability while Suzy's throw hits it with a certain, possibly different, probability. They both throw, and Suzy's stone hits the bottle, and breaks it. In that case we would say

[6] Unlike in other dependency theories discussed above, in probabilistic dependency theories there can be a causal relation (between kinds) even when the effect does not occur (at the individual level), since all that is needed to calculate a causal relation is the probability of the effect's occurring under certain conditions. As we will see in section 2.3.1, this property could be useful in understanding cases of defeasible causation in language, such as non-culmination of accomplishments.

Suzy's throw was the cause of the bottle breaking—and indeed her throw raised the probability of the bottle's breaking. However, Billy's throw also raised the probability of the bottle's breaking, although his throw was not the cause (Hitchcock 2010).

Additionally, probability raising is apparently not *necessary* for a factor to be considered a cause; cases exist where C is a cause of E but C does not raise the probability of E. Imagine that Suzy throws her rock with a 25% chance of shattering of the bottle. If Suzy had not thrown the rock, Billy would have done so, with a 70% chance of shattering. In this example, Suzy's throw would be the cause of the shattering, even though it lowered the chance of that effect (from 70% to 25%) (Hitchcock 2010).[7]

As usual, problems such as these are probably not insurmountable, but any viable solution would be expected to bring complications to the theory. Probabilistic dependency theorists have addressed such problems by getting more specific about the background contexts on which probabilities are calculated (Cartwright 1979; Skyrms 1980), as well as by recognizing differences between singular and general (kind) causation (Eells 1991; Hitchcock 2004), since probabilities can arguably only be calculated for kinds of events, not for individual events.

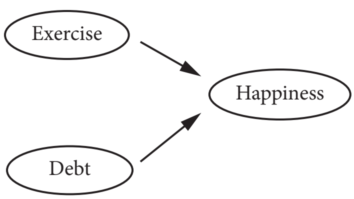
2.2.1.4 Causal modeling approaches to causation One particular formal implementation of the dependency view of causation has had a wide-ranging influence on a number of fields. As Williamson (2009) points out, the formalism of Bayesian networks developed in the 1980s (Pearl 1988; Neapolitan 1990) provided an efficient way to think about causal connections at a time when causal explanations were out of fashion in scientific fields, in part due to Russell's (1913) attack on the notion of causation as being unnecessary for scientific explanation.

A causal Bayesian network represents the causal structure of a domain and its underlying probability distribution. The causal structure of the domain is represented by a directed acyclic graph of nodes and arrows, whereas the probability distribution consists of the conditional and unconditional probabilities associated with each node. The alignment of these two kinds of information allows us to make predictions about causal relationships using probability theory. A simple causal Bayesian network is shown in Fig. 2.1. Each node in the network is associated with an unconditional, prior probability. For example, in the the network shown in

[7] Another example of how probability raising is not necessary for causation is seen in cases where the influence of one cause is overwhelmed by the influence of another (Cartwright 1979; Hitchcock 2010). For example, under the right circumstances, the probability of cancer might be less in the presence of smoking than the probability of cancer in the absence of smoking, that is P(cancer | smoking) < P(cancer | not smoking). Clearly, smoking causes cancer, but a positive correlation between cancer and smoking might be masked, or even reversed in the presence of another cause. Imagine a situation in which not smoking is correlated with living in a city, breathing highly carcinogenic air. In such a situation, not smoking could be more strongly associated with cancer than smoking, but the causal relationship between smoking and cancer could remain. Such reversals are widely known as examples of Simpson's paradox; see Kistler (Ch. 4, this volume) for additional discussion.

Fig. 2.1, exercise is associated with a 0.5 probability of being true and a 0.5 probability of being false, while debt is associated with a 0.2 probability of being true and a 0.8 probability of being false. The arrows in this graph represent causal relations (in the broad sense). In Fig. 2.1, the arrows from exercise and debt to happiness convey that these two variables affect happiness. The exact way in which they do so is described in the probability table associated with happiness, which specifies several conditional probabilities: for example, the probability of happiness being present when one exercises but also has debt, i.e. P(Happinesss | Exercise and Debt), is 0.6, and the probability of not being happy when one exercises and has debt, i.e. P(∼Happinesss | Exercise and Debt), is 0.4. The conditional probabilities specified in the probability table specify that exercise raises the probability of happiness, whereas debt lowers the probability of happiness. It is in this manner that a causal Bayesian network can represent both facilitative and inhibitory causal relations, and it is for this reason that the arrows are causal in a broader sense than is encoded in the meaning of the verb *cause*. Roughly, the arrows mean something like *influence* or *affect*.

Exercise	
T	F
0.5	0.5



Exercise	Debt	Happiness T	Happiness F
T	T	0.6	0.4
T	F	0.9	0.1
F	T	0.3	0.7
F	F	0.5	0.5

Debt	
T	F
0.2	0.8

Fig. 2.1 Causal Bayesian network with associated probability tables

Causal Bayesian networks allow us to reason about causation in more than one dimension, i.e. in networks rather than mere chains. For example, they allow us to predict—using Bayes' rule—the probability of certain variables being true when other variables are either true or false, both in the direction of causation and diagnostically, i.e. working from effects to causes. However, in order to understand where the causal arrows themselves come from—i.e. when we are justified in asserting a causal relation between two variables—more must be said.

In order for a Bayesian network to qualify as a causal Bayesian network, it has been argued (Hausman and Woodward 1999) that its probabilistic dependencies and arrows must honor the *Causal Markov Condition*.[8] The Causal Markov Condition holds that a variable C will be independent of every variable in a network except its effects (i.e. descendants) (e.g. E), conditional on its parents. Hausman and Woodward (1999) use the Causal Markov Condition as part of a sufficiency condition on causation: if C and E are probabilistically dependent, conditional upon the set of all the direct parents of C in the given Causal Markov Condition-satisifying model, then C causes E. However, late pre-emption provides a counterexample to this sufficiency condition (i.e. a case where the condition holds but the intuition is that C does not cause E).

Another way to characterize causation in a causal Bayesian network is in terms of the notion of *intervention* (Pearl 2000; Woodward 2009). An intervention is a process by which a variable in a network is set to a particular value. The notion of intervention is closely related to our sense of causation. In effect, interventions allow us to conduct counterfactual reasoning. If C causes E, then intervening on, or "wiggling" the value of C should result in corresponding changes in the value of E. If we can intervene to counterfactually change the value of C to any possible value, and can still predict the probability of the value of E being true, we can be confident that C causes E. For example, suppose that you want to find out if a switch being in an up position causes a light to be on. The natural thing to do is to try the switch and see if the light is on when the switch is up and off when the switch is down. If the status of the light depends on the switch position in all positions (i.e. on and off), we feel justified in concluding that setting the switch to the up position causes the light to turn on. Note that wiggling the value of E should have no effect on the value of C. Thus, interventions allow us to determine the direction of the causal arrow. Interventions thus provide us with an alternative sufficient condition on causation: if interventions on C are associated with changes in E, C causes E (Hausman and Woodward 1999).[9]

[8] Strictly speaking, dependence means $P(E \mid C) \leftrightarrow P(E)$, so an arrow will be justified only if this is true. For example, if $P(E \mid C) = .5$ but $P(E) = .5$, there will be no arrow between C and E in a model that satisfies the Causal Markov Condition.

[9] The notion of intervention may seem to approach the notion of agency, and indeed an alternative approach to causation has pursued this idea. Woodward (2009) separates 'manipulation-based' accounts into interventionist theories such as we have described, which refer merely to intervention by whatever external cause, and agency theories (e.g. Menzies and Price 1993), which define causation in terms of explicitly animate or human agency. Menzies and Price propose that rooting the theory in our personal experience of agency keeps the theory from being circular, which is a desirable outcome (and foreshadows the production class of theories, section 2.1.2). On the other hand, there are counterexamples to their claim that agency is a sufficient condition for causation, including cases where there is no possibility for an agentive manipulation (Hausman and Woodward 1999); at any rate, an animate intervener is not necessary in order to define an intervention, just as inanimate entities can be causers (see section 2.2.2).

The light-switch example suggests an additional way to use causal models: it is possible to model deterministic causal structures as well as probabilistic causal structures. The special case where the values of each variable are limited to 1 and 0 (true and false, on and off) yields tables reminiscent of familiar Boolean truth tables and is therefore possibly of more interest to linguists (though of less interest to probability theorists; Bayes's rule is no longer relevant). One example of such an approach is Hitchcock (2010), which shows how a deterministic model accounts for the problem of late pre-emption that so bedeviled previous dependency theories of causation.

In addition to the node-and-arrow notation, Hitchcock presents his causal networks in terms of *structural equations* (see also Sloman et al. 2009).[10] Consider the causal network in Fig. 2.2.

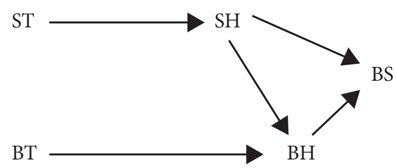


FIG. 2.2 An example of late pre-emption in terms of a direct graph (Hitchcock 2010)

The causal graph shown in Fig. 2.2 specifies the late pre-emption scenario discussed earlier, in which ST corresponds to Suzy's throw, SH to Suzy's ball hitting the bottle, BS to the ball's shattering, BT to Billy's throw, and BH to Billy's ball hits the bottle. The causal network shown in Figure 2.2 can be re-expressed in terms of structural equations as follows. The ":=" relation is an asymmetrical relation, read as "gets", in the opposite direction of the arrows.[11]

(1) SH := ST
BH := BT
BS := SH v BH
BH := BT and ∼SH

[10] These are equally available for the probabilistic case; here we examine the special case where values are either 1 or 0. One advantage to this notation over the node-and-arrow plus table notation is that it allows us to see at a glance whether a value of a parent variable has a positive or negative effect on the probability of a certain value of the child variable (since, as we have noted, both positive and negative effects are represented with the same kind of arrow).

[11] Note that there is an arrow from BT to BH and from BH to BS, even though we do not want to say that BT (or BH) causes BS. These arrows and the associated truth value distributions satisfy the Causal Markov Condition (see section 2.2.1.4), however. This failure in the face of late pre-emption shows that the Causal Markov Condition alone is not the correct sufficiency condition for causation.

As Hitchcock notes, an interventionist approach can offer a solution to the problem of late pre-emption. Given the equations in (1), we can simulate different scenarios by setting the variables to different values. For example, we could simulate the late pre-emption scenario by setting Suzy's hitting the bottle, SH, and Bill's hitting the bottle, BH, to "1". When this is done, the value of BS would be 1 as well; the bottle would shatter.

As we have said, a variable C is taken to cause E in a certain scenario if the values of E co-vary when C is wiggled and all external variables are held constant *at their actual values*[12] in that scenario. In the Suzy and Billy late pre-emption case, the actual values of Suzy's and Billy's throw, ST and BT, would be 1, the actual value of Suzy's hit, SH, would be 1, while the actual value of Billy's hit, BH, would be 0. It is interesting to see how such a graph is able to account for the intuition that, when both Suzy and Billy throw their rocks, with Suzy throwing first, we would describe Suzy and not Billy as the cause. To test whether Suzy is the cause, we need to hold BH fixed. In the actual scenario, BH is 0. Under these conditions, the value of BS would covary with the value of ST, implying that Suzy's throw is a cause of the shattering. To test whether Billy's throw is a cause, we need to hold SH to the value it has in the actual scenario, i.e., 1. With SH set to 1, the bottle would shatter regardless of the value of BT and the counterfactual test for BT would often be incorrect, offering evidence against BT being the cause of the shattering.

It is worth emphasizing the reason why the structural equation approach to encoding counterfactuals is able to account for late pre-emption. The reason why it succeeds is due to the asymmetry in the values of SH and BH. These two variables take on different values because of the requirement to freeze values at only their actual values; SH can be set to 1 while BH is set to 0, while the converse is not possible.

Causal Bayesian networks and structural equation modeling have several attractive properties: not only do they allow us to go beyond simple causal chains to specify causal networks in which some nodes have more than one parent (especially useful for counterfactuals; e.g. Dehghani et al. 2012), but they can be used to model both probabilistic and deterministic causation. Furthermore, they suggest straightforward accounts for late pre-emption. However, some concerns linger.

There are cases in which both of the sufficiency conditions mentioned hold between C and E, but C does not cause E: suppose that a villain gives the king poison (C), which causes the king's adviser to give the king an antidote, which on its own would kill the king but which neutralizes the poison harmlessly so the king survives (E) (Hitchcock 2007). In that case, it turns out that the intervention condition predicts C to be a cause of E, but we do not have the intuition that

[12] This requirement is an analogue to Lewis's similarity metric over possible worlds, relating them to the actual world: in both cases, certain other potentially interfering variables must be held constant at their actual value in order to determine if C causes E.

C causes E. Of course, one might propose a different sufficiency condition, and/or additional constraints on the model to explain these facts.

A more serious issue is the question of what these models are for. As Hausman and Woodward (1999) point out, it is curious to characterize causation in terms of intervention, which is itself arguably a causal notion. Such a characterization of causation is uninformative at best and circular at worst. This is not a problem if the models are used to analyze structures in which the direct causal relations are already known, and the question at hand is to find out how certain direct causal relations combine to yield causal relations in a complex structure. However, if these models are meant to be a theory of causation, and intervention is disqualified for circularity, it is only the Causal Markov Condition and other such conditions on the models that bear on the question of what causation actually is (and as demonstrated in the case of late pre-emption, the Causal Markov Condition is not enough to guarantee causation, though other conditions on the model can and have been added; see e.g. Woodward 2009). This is fine, but the complexity of the Causal Markov Condition and whichever additional conditions would be added to it raises the question of whether these are merely tests for whether certain structures can arise from causation, rather than accounts of our intuitive notion of causation itself (Mackie 1974).

2.2.2 Production theories

In the previous section, we touched on the major kinds of dependency theory of causation: logical dependency, counterfactual dependency, probabilistic accounts, and Bayesian and causal modeling accounts. What they have in common is the idea that causation can be explained by means of a dependency between the cause and the effect. The hope that motivates dependency theorists is that causation can be reduced to correlation or regularity if the conditions are pruned and the potentially confounding variables are fixed correctly. As we have seen, this hope is in large part justified by the success that such theories have had in providing appropriate sufficiency conditions for causal intuitions.

On the other hand, we seem also to have an intuition that something more than correlation or regularity is involved in causation (Pinker 2008; Saxe and Carey 2006). Hume recognized as much. He acknowledged that we often associate causation with a sense of force and energy. But for Hume, these were mental experiences that accompanied causation. He maintained that these notions could not be the basis for our understanding of causation, on the assumption that they could not be objectively observed. For Hume, these notions were imposed on experience by the mind, rather than experience imposing these notions on the mind. Ideas of force or energy are epiphenomena of our personal, subjective interactions with causation (Fales 1990; White 1999; 2006; 2009; Wolff and Shepard 2013).

It may be, however, that ideas of force and energy are more central to the notion of causation than was recognized by Hume, or for that matter by dependency theories, which are largely descendants of the Humean perspective. One argument for why force and energy may be central to the notion of causation emerges when we consider the range of properties commonly associated with causal relationships.

One such property is temporal order: if C causes E, C must precede, or at least be simultaneous with E (Lagnado et al. 2007). This temporal relation between cause and effect is thus a necessary condition for causation. Like correlation, this relation is clearly not a sufficient condition for causation. Nonetheless, temporal precedence has been shown to be a stronger indication that C causes E than even correlation between C and E (Lagnado and Sloman 2006). The relationship between causation and temporality has been discussed by some linguistic researchers as well (e.g. Shibatani 1973a; Talmy 1976), though it has been ignored in much of the literature on the syntax–semantics interface.[13]

A second property is having a physical link between cause and effect (Salmon 1984; Walsh and Sloman 2011). This property requires some qualification. By physical link, we do not necessarily mean a direct physical contact; rather, that the cause and effect are linked in some way either directly or indirectly, through a chain of physical connections. This property appears not to be a necessary condition of causation, because of a large class of exceptions to this property that rely on "spooky action at a distance" (Einstein's famous description of quantum entanglement). This class includes not only gravity, electromagnetism, and quantum entanglement, but also magic and divine intervention.[14] Exactly when there is no plausible physical link, spooky influences such as these are called upon to justify impressions of causation.

These properties are problematic for dependency theories because these theories do not provide motivation for why these properties are relevant to causation. Temporal precedence or simultaneity, for example, is handled by stipulation, i.e. it needs to be explicitly stated in all these theories that the cause precedes the effect or occurs at the same times as the effect (Wolff, Ch. 5, this volume). The physical-link property is rarely if ever mentioned by dependency theorists. Why are these properties associated with causation? And is an answer to this question crucial to our notion of causation?

Our personal view is that the answer to that question is important, and since dependency approaches to causation give us no understanding of why these properties are relevant to causation, we must look elsewhere for an answer. In theories of causation based on concepts of force or energy, these properties of causation fall out naturally. A force is exerted or energy is transmitted, before or simultaneously with the effect that is provoked. Most forces also require a physical link, except, notably, for the class of spooky influences. These facts suggest that concepts such as

[13] See Copley and Harley (to appear) for a recent linguistic discussion of the difference between launching causation, in which the cause precedes the effect, and entrainment, where the cause and effect happen at roughly the same time (Michotte 1946/1963).

[14] Chains involving "social forces" might be thought to be part of this class, but as long as there is transmission of information from one person to another, there is still a physical link.

force and energy provide a necessary part of our notion of causation, and that Hume had it exactly backwards: that force and energy are in fact the basis of our notion of causation, while correlation and regularity are the epiphenomena.

Theories that characterize causation in terms of concepts such as force and energy view causation as a *production* or process. The production may involve a transmission of conserved quantities such as energy (Dowe 2000; Kistler 2006a; Ch. 4, this volume). It may also be viewed in terms of causal powers, namely the ability of entities to transmit or receive a conserved quantity (Mumford and Anjum 2011a). Yet another approach would be in terms of forces being imparted, for instance, by an agent to a patient (as in the parallel cognitive linguistic tradition, e.g. Talmy 1988; 2000; Gärdenfors 2000; Warglien et al. 2012; Croft 1991; 2012; Ilić, Ch. 7, this volume; also Wolff 2007; Ch. 5, this volume). See also Copley and Harley (Ch. 6, this volume) for a more abstract view of forces.

Theories of causation that characterize causation in terms of transmission include Salmon's (1984; 1998) mark transmission theory. In this theory, causation is understood primarily as a process rather than as a relation between events. A causal process is understood as a transmission of a causal mark, i.e. a propagation of a local modification in structure. A causal process would be instantiated if, for example, one put a red piece of glass in front of a light. In such a case, the red glass would impart a mark on a process that would transmit the mark to a different location, such as a wall.

Salmon's theory's greatest strength may be in its ability to distinguish, in certain circumstances, causation from pseudo-causation. However, because the theory emphasizes processes over events, it does not provide a direct definition of what counts as causation. It is not hard to imagine how Salmon's theory might be extended to provide such a definition. To say that A causes B might be to say, in effect, that a mark is propagated from A to B. In some cases, a procedure can be specified for determining whether a mark has been propagated. In the case of the light filter, one can check to see what happens when the filter is removed. However, in many other cases, procedures for determining whether a mark has been propagated are less clear. For example, in the ordinary billiard-ball scenario, what is the mark and how do we know it has been propagated? If the procedures cannot be specified, then the legitimacy of the causal relation should be ambiguous; but in the case of billiard-ball scenarios, at least, the the legitimacy of the causal relation is not in doubt. It might be possible, through further elaboration of the theory, to address this challenge. In particular, in order to make the criteria for causation easier to assess it would help to have a clearer idea of the notion of a mark.

A potential solution to this problem is offered by Kistler (2006a), who proposes a transmission theory of causation that brings back the idea of causation being a relation between a cause and an effect. According to this theory, "Two events c and e are related as cause and effect if and only if there is at least one conserved quantity P, subject to a conservation law and exemplified in c and e, a determinate amount of which is transferred between c and e." Kistler (2006a) goes on to define "transference" as present if and only if an amount A is present in both events. In order for this to occur, events c and e must be located in space and time in such a manner that allows for the transference. In particular, the transference process requires spatial and temporal contiguity and implies that causation must take place over time (but does not, according to Kistler (2006a), necessarily imply that the cause precedes the effect).

Kistler's (2006a) proposal that causation involves a transference of a conserved quantity builds on a highly influential theory by Dowe (2000). According to Dowe's Conserved Quantity Theory, there are two main types of causation: persistence (e.g., inertia causing a spacecraft to move through space) and interactions (e.g., the collision of billiard balls causing each ball to change direction). Causal interactions are said to occur when the trajectories of two objects intersect and there is an exchange of conserved quantities (e.g. an exchange of momentum when two billiard balls collide). Unlike earlier theories, exchanges are not limited to a single direction (i.e., from cause to effect). One problem that has been raised for Dowe's theory—and that also applies to transference theories—is that such a theory is unable to explain the acceptability of a number of causal claims in which there is no physical connection between the cause and the effect. In particular, such a theory seems unable to handle claims about causing preventions or causation by omission (Schaffer 2000; 2004). Consider, for example, the preventative causal claim, "Bill prevented the car from hitting Rosy", assuming a situation in which Bill pulls Rosy out of the way of a speeding car. Such a causal claim is acceptable, even though there was no physical interaction between Bill and the car. Perhaps even more problematic are causal relations resulting from omissions, as when we say: "Lack of water caused the plant to die." The acceptability of such a statement cannot be explained by transmission or interaction theories since, plainly, there can be no transmission of conserved quantities from an absence.

Another type of production theory holds that causation is specified in terms of forces (Copley and Harley, Ch. 6, this volume; Talmy 1988; 2000; Gärdenfors 2000; Croft 1991; 2012; Ilić, Ch. 7, this volume; Warglien et al. 2012; Wolff 2007; Ch. 5, this volume). One such theory is Wolff and colleagues' force dynamic model (Wolff 2007; Wolff et al. 2010). According to this model, causation is specified in terms of configurations of forces that are evaluated with respect to an endstate vector. Different configurations of forces are defined with respect to the patient's tendency towards the end-state, the concordance of agent's and patient's vectors, and the resultant force acting on the patient. These different configurations of forces allow for different categories of causal relations, including the categories of cause-and-prevent relations. In a preventative relationship, there is a force acting on the patient that pushes it towards an end-state, but the patient is then pushed away from the end-state by the force exerted on it by the agent. Philosophers and cognitive scientists have argued that transmission theories are unable to explain preventative relationships, as well as the notion of causation by omission. In order to capture these phenomena, it has been argued that theories of causation must go beyond a production view of causation to include, perhaps, counterfactual criteria for causation (e.g. Schaffer 2000; Dowe 2001; Woodward 2007; see also Walsh and Sloman 2011). Interestingly, Talmy's theory of force dynamics and, relatedly, Wolff's dynamics model are able to explain how the notion of prevention can be specified within a production view perspective without having to incorporate distinctions from dependency theories, such as counterfactual criteria. Wolff et al. (2010; see also Wolff, Ch. 5, this volume) also show how a production view of causation is able to handle the phenomenon of causation by omission—a type of causation which, according to several philosophers, is beyond the explanatory scope of production theories (Schaffer 2000; Woodward 2007).

Production theories have several attractive qualities. As already noted, they motivate why the concept of causation is associated with temporal and spatial properties. They also provide relatively simple accounts of people's intuitions about scenarios that are problematic for dependency theories, such as late pre-emption. In the case of late pre-emption, there are two possible causers and one effect. For example, in the Suzy and Billy scenario, Suzy and Billy both throw rocks at a bottle and the bottle breaks, but Suzy's rock hits the bottle first. Intuition says that Suzy's throw caused the bottle to break. This intuition falls out naturally from production theories: in transmission theories, in particular, Suzy is the cause of the breaking because it was from Suzy's rock that conserved quantities were transmitted to the bottle, while in force and power theories, Suzy's throw is the cause because it was from Suzy's rock, not Billy's, that force was imparted upon the bottle.[15]

Though production theories have several strengths, they also face several significant challenges. Without further qualification, production theories require that knowing that two objects are causally related entails being able to track the transmission of conserved quantities linking two objects. Such a requirement often does not hold for a wide range of causal relations. For example, common sense tells us that there is a causal relationship between the light switch and the lights in a room, but most of us could not say exactly how conserved quantities are transmitted through this system. As argued by Keil and his colleagues (Rozenblit and Keil 2002; Mills and Keil 2004), people often feel as if they understand how everyday objects operate, but when they are asked to specify these operations, it becomes clear that they have little knowledge of the underlying mechanisms. Keil and his colleagues refer to this phenomenon as the "illusion of explanatory depth". The

[15] As we have seen, Lewis in his later work (e.g. 2000) responds to criticism of his counterfactual approach by proposing that causation must be evaluated not just on whether C and E are true, but on finer properties of the events referred to by C and E. Late pre-emption is accounted for by noting that counterfactually changing the properties of Sally's throw changes the properties of the bottle's breaking, while changing the properties of Billy's throw does not. This interest in finer properties of events, rather than just truth values, is perhaps the closest approach of a dependency theory to the spirit of production theories, though it should be noted that Lewis's 2000 theory still relies on the notion of change rather than the notion of energy.

illusion of explanatory depth presents production theories with a challenge: how can causal relations be asserted of situations in which the underlying mechanism is not known? Production theory advocates might appeal to people's general knowledge of how things are likely physically connected, but such a move introduces uncertainty into their knowledge of causal relations, and if there is uncertainty, why not simply represent the causal relations in terms of probabilities and relationships between probabilities? Currently, there are no simple answers to such a challenge (but see Wolff and Shepard 2013).

A second major challenge for production theories concerns the problem of how such an approach might be extended to represent abstract causal relations. Production theories are clearly well suited for causal relations in which the quantities being transmitted are grounded in the physical world. For example, production theories seem especially well suited for explaining the acceptability of statements such as *Flood waters caused the levees to break*, or *The sun caused the ice to melt*. Much less clear is how a production theory might represent statements such as *Tax cuts cause economic growth*, or *Emotional insecurity causes inattention*. Obviously, such abstract instances of causation cannot be specified in terms of physical quantities, so how might they be represented? According to some theorists, abstract causation might be represented in a fundamentally different manner than concrete causation. Such a view has been dubbed *causal pluralism* (Psillos 2008; see Kistler, Ch. 4, Wolff, Ch. 5, this volume). Another possibility is that abstract causation might be represented in a manner analogous to physical causation (Lakoff and Johnson 1999; Wolff 2007). While explaining abstract causation in terms of metaphor might be an easy move to make, questions soon arise about such an explanation's testability. As with the problem created by the illusion of explanatory depth, there is currently no simple answer to the challenge raised by abstract causation for production theories (but see Wolff, Ch. 5, this volume, for an attempt).

2.3 Linguistic phenomena to which causation is relevant

Recall that our main purpose in this chapter was to demonstrate that theories of causation are relevant to linguistic theory and vice versa. Having presented the state of research on causation as we see it, we now turn to examine three linguistic domains to which causation is relevant. We will argue that a more sophisticated understanding of different theories of causation has real potential to advance our knowledge of these phenomena, and conversely, that these linguistic analyses, especially those concerning data from less familiar languages, should be taken seriously by philosophers and cognitive scientists working on causation.

The phenomena we will examine are:

Defeasibility: A causal relation has been proposed between a cause and effect in the two sub-events of accomplishments, but in certain environments, such as progressives, non-culminating accomplishments, and frustratives, the effect does not occur.

Volition, intention, and agency: Numerous linguistic phenomena seem to distinguish animate agents from inanimate causers. Language thus appears to be sensitive to whether the causing entity is volitional (or intentional) or not, though volitionality seems to not always be quite the right notion; rather, a broader causal notion that subsumes but is not limited to volitionality is called for. We argue that disposition fits the bill.

Representations of causal chains: How conceptual representations of events and participants in causal chains are mapped onto language, both to syntactic event chains and within certain lexical items.

Apart from being of interest in themselves, discussion of these phenomena will allow us to demonstrate several ways in which the choice of causal theory can bear on linguistic theory. For example, one instance of defeasible causation—non-culminating accomplishments—illustrates the heuristic that complex semantics should be adduced only when there is visible morphology in some language. Since different theories of causation predict different parts of causal semantics to be complex, choosing a causal theory whose distribution of complexity matches what is seen in the morphosyntax has the opportunity to greatly simplify the (morpho)syntax–semantics interface. Another example of how causal theory can be of use to linguists is when grammatical evidence suggests that certain concepts are linked, for example volitionality and the ability of inanimate objects to be the external arguments of activities (Folli and Harley 2008). In such cases, using a causal theory that explains the link between these concepts is preferable to one that does not, since it has a better chance of informing the semantic theory. Finally, it is fairly easy to see how causal theory can bear on the question of how conceptual causal chains correspond to the causal chains that are represented in language. In the mapping from conceptual causal structure to semantics, certain phenomena corresponding to components of causal theories are observed at the syntax–semantics interface, and/or the lexicon, while others are not. Ideally, for whichever causal theory is chosen, the components that are observed linguistically should be those that are important to the theory; conversely, nothing important to the causal theory should be completely invisible to language.

Two caveats must be mentioned. First, we are prepared for the possibility that different causal theories may be useful for different linguistic phenomena. This possibility is related to the notion of causal pluralism discussed earlier. However (and this is the second caveat), it should be clear that the study of language's relation to a mental representation of causation by definition has only to do with how causation is represented in the mind, not with the actual nature or metaphysics of causation. To the extent that certain philosophers are concerned with the metaphysics of causation rather than the mental representation of causation, a causal pluralism in language would not bear directly on their findings, although the use