

Situations and Syntactic Structures

Situations and Syntactic Structures
Rethinking Auxiliaries and Order in English

<optional edition>

Gillian Ramchand

The MIT Press
Cambridge, Massachusetts
London, England

© 2016 Massachusetts Institute of Technology

All rights reserved. No part of this book may be reproduced in any form or by any electronic or mechanical means (including photocopying, recording, or information storage and retrieval) without permission in writing from the publisher.

For information about special quantity discounts, please email special_sales@mitpress.mit.edu.

This book was set in Syntax and Times Roman by the author.

Printed and bound in the United States of America.

Library of Congress Cataloging-in-Publication Data

<Situations and Structure> / <Gillian Ramchandr> .
p. cm.

Includes bibliographical references and index.
ISBN ????

10 9 8 7 6 5 4 3 2 1

Contents

| | | |
|--------------|--|-----------|
| Preface | xi | |
| 1 | Introduction to Events and Situations in Grammar | 1 |
| 1.1 | Linguistic Generalizations and Constraints on the Syn-Sem Mapping | 3 |
| 1.2 | Event Properties and Event Instantiations | 8 |
| 1.2.1 | Philosophical Antecedents | 9 |
| 1.2.2 | A Quotational Semantics for Natural Language | 15 |
| 1.3 | The Grammar of Auxiliation | 22 |
| 1.3.1 | Ordering | 23 |
| 1.3.2 | Lexical Specification and Polysemy | 24 |
| 1.4 | Morphology and Spanning | 28 |
| 1.5 | Roadmap | 38 |
| 2 | The Progressive in English | 41 |
| 2.1 | The Syntax of the Progressive: The Progressive is in the First Phase | 42 |
| 2.1.1 | Expletive Associates | 43 |
| 2.1.2 | VP fronting and pseudoclefts | 44 |
| 2.1.3 | British nonfinite <i>do</i> -substitution | 45 |
| 2.2 | The Semantics of the Progressive | 49 |
| 2.3 | The Present Proposal | 59 |
| 2.3.1 | The Progressive as an ‘Identifying State’ | 60 |
| 2.3.2 | Relating Event Properties to Situational Particulars | 64 |
| 2.4 | Conclusion | 72 |
| 3 | The Passive and its Participle | 75 |
| 3.1 | The Stative Participle in <i>-en/ed</i> | 77 |
| 3.1.1 | Event Implications | 78 |
| 3.1.2 | An Implementation in Terms of Reduced Spans | 84 |
| 3.2 | The Eventive Passive | 95 |
| 3.2.1 | Blocking | 98 |
| 3.2.2 | Interaction of Passive and Progressive within EvtP | 104 |

| | | |
|-------|---|-----|
| 3.2.3 | Adjectivalization | 107 |
| 4 | The Participle, the Perfect and the Spatiotemporal Domain | 113 |
| 4.1 | Semantic Background to the English Perfect | 116 |
| 4.1.1 | Aktionsart Sensitivity | 119 |
| 4.1.2 | Temporal Modification and the Present Perfect Puzzle | 121 |
| 4.1.3 | Lifetime Effects and Current Relevance | 122 |
| 4.2 | The Proposal | 124 |
| 4.2.1 | Accounting for the Pragmatic/Lifetime Effects | 129 |
| 4.2.2 | Temporal Modification and the Perfect | 133 |
| 4.3 | Summary | 137 |
| 5 | Modals and the Spatiotemporal Domain | 145 |
| 5.1 | Introduction | 145 |
| 5.2 | What We Know: Syntax | 147 |
| 5.2.1 | Thematic Relations: 'Raising' vs. 'Control' | 151 |
| 5.2.2 | Scope wrt to a Quantified Subject | 152 |
| 5.2.3 | Modals and Symmetric Predicates | 153 |
| 5.2.4 | Ordering and Typology | 155 |
| 5.2.5 | Taking Stock | 158 |
| 5.3 | Interaction with Negation | 159 |
| 5.3.1 | English Modals are Idiosyncratic | 160 |
| 5.3.2 | A Semantic Account of the Idiosyncrasy? | 163 |
| 5.4 | Semantic Properties of Modals | 167 |
| 5.4.1 | The Classical Theory | 167 |
| 5.5 | Circumstantial Modals as Modifiers of Spatiotemporal Properties | 174 |
| 5.6 | Conclusion | 183 |
| 6 | Modals and Generalized Anchoring | 185 |
| 6.1 | Temporal Properties of Modals | 186 |
| 6.2 | Epistemic Modals as Modifiers of Assertions | 196 |
| 6.2.1 | Aktionsart Sensitivity of Epistemic Modals | 199 |
| 6.2.2 | Modals Embedding the Perfect | 203 |
| 6.3 | Evidence for Choice among LIVE-ALTERNATIVES as Basic to Modal Meaning | 206 |
| 6.3.1 | Relation to Dynamic Modal Interpretations | 206 |
| 6.3.2 | The Semantic 'weakness' of Universal Modals. | 207 |
| 6.3.3 | The Puzzle of Deontic Modality and its interaction with Disjunction. (The Paradox of Free Choice Disjunction (Ross 1941)) | 209 |
| 6.4 | Conclusion | 210 |
| 7 | Summary and Future Prospects | 215 |
| 7.1 | Architecture and Semantic Zones | 215 |
| 7.2 | Insertion: Lexical vs. Functional Items | 219 |

| | | |
|--------------|--|-----|
| 7.3 | Summary of the Pieces | 221 |
| 7.4 | Auxiliary Ordering Revisited | 226 |
| 7.5 | Open Questions and Further Research | 228 |
| 7.5.1 | The Nominal Domain | 228 |
| 7.5.2 | On the Universal vs. the Language Specific | 229 |
| 7.5.3 | The Future and the Search for Explanations | 230 |
| | Bibliography | 231 |

Preface

In some ways I have been writing this book in my head for twenty years, working on various aspects of the syntax and semantics of the verbal extended projection in English and other languages. The immediate impetus came from the collaborative paper I wrote with Peter Svenonius on reconciling minimalist and cartographic approaches to phrase structure (Ramchand and Svenonius 2014), and I knew that the compositional semantics component of that paper was a huge promissory note that had to be redeemed if the enterprise were to succeed. The present book is motivated by a conviction about what the relationship between phrase structure and semantic interpretation should look like, a conviction I found that many shared, but which was rather difficult to actually implement. Implementing the intuition required some radical changes to the assumed semantic ontology for natural language, ones which I believe are more in line with internalist intuitions (cf. Chomsky 1995, Pietroski 2005), without giving up on a system that grounds interpretation in truth. The formal semantic framework I have in mind to underpin the kind of system I propose is a version of Kit Fine's truthmaker semantics (Fine 2014), using situations as exact verifiers for natural language clauses. The system is quite different from the kind of semantics that takes worlds as its foundation, and in which, instead, possible and impossible situations are *primitives* of the ontology. The intuition that is important to me is that the syntax of natural languages gives evidence on the meaning side for a natural language ontology which might be quite different from the one that seems most compelling from a purely metaphysical point of view. Like Fine, (and Moltmann 2018) I am more concerned with *how* language puts meaning together than with how truthmakers are connected to a metaphysics of the real world (if indeed such a thing is even possible).

One crucial innovation in the system requires reifying the linguistic symbol itself as an object in the ontology. I do not think this is a 'trick', but in fact simply acknowledges a feature of the natural language system that is very important— self referentiality and a deep indexicality (relativization to

the particular speaker) which I believe has desirable consequences that extend far beyond the scope of the present monograph.

The first few chapters of this book were written while I was on sabbatical at the University of Edinburgh in the second half of 2015. I thank the University of Tromsø for its generous sabbatical provision, and its support of pure theoretical research, and the University of Edinburgh for being a welcoming host. In the early stages of writing I benefitted a great deal from correspondence with Lucas Champollion, and to discussions with Ronnie Cann to whom I am very grateful. I would also like to thank audiences at the LOT Winter School in Nijmegen in January 2017 and to audiences at a mini course at the University of Budapest in February 2017. I would like to thank Marcel den Dikken and Eva Dekány for inviting me to teach at the latter event. I also benefitted from discussion and feedback at Daniel Altshuler's UMass semantics seminar in Spring 2017. I would further like to thank Daniel Altshuler personally and Miriam Butt for very useful detailed comments on an intermediate draft, and Sergey Minor for invaluable and detailed feedback on the whole prefinal manuscript. Special thanks go to Robert Henderson for turning up in Tromsø in late 2016 and giving a talk at our colloquium series on ideophones, which contributed the final piece of the puzzle. I remember that eureka moment very well when I realised that the exoticism of ideophones was just 'the truth standing on its head to get attention'. A big thank you goes to my colleagues at the University of Tromsø and in particular the CASTLFish milieu (Tarald Taraldsen, Antonio Fábregas, Sergey Minor, Peter Svenonius and Björn Lundquist deserving special mention in this regard), for reacting and commenting on all things related to syntax and morphology, for providing the intellectual frame for the kinds of questions I find myself asking, and for providing standards for the kinds of answers that satisfy.

Finally, even though this book did not benefit from any direct or detailed discussions with Angelika Kratzer, she is in many ways the forerunner and inspiration for the research agenda here. Her work on the syntax-semantics interface is interwoven with these pages, and this particular book would have been impossible without her research as backdrop and standard.

I thank Anne Mark and the team at MIT Press for their engagement and professionalism in bringing this work to publication.

*Gillian Ramchand
October, 2017*

1

Introduction to Events and Situations in Grammar

The primary reason for the use of events in the semantics of natural language is empirical. Starting with Davidson (1967), there is by now a large body of evidence that they are necessary ingredients in the most empirically adequate descriptions of the way language works. The obvious application is in the semantics of verbs themselves, a class of words found in every natural language we know of (see Baker 2003). What has been more controversial, and where there are still open and lively debates, is the exact nature of the interpretational ontology, and how it connects to the compositional semantics of natural language.

My purpose in this book is not to argue for the existence of verbal events per se, or particular details concerning their nature and internal structure (I refer the reader to Truswell, to appear, for the state of the art). It *is* however, a book that relates directly to issues of semantic ontology, and the way we set up our compositional semantics so as to be properly integrated with robust facts about the syntax and morphology of natural language.

Verbal meanings are remarkably diverse, albeit within certain constrained abstract limits.¹ At a very basic level, we need a place-holder variable as the unity to which the different core properties of an event description can be ascribed. This is what I take to be the fundamental insight of Davidson (1967). We also need, at the end of the day, to be able to construct arbitrarily complex coherent and unified descriptions of the world, and assert their existence. These coherent unified situations are built up cumulatively from a combination of the verb and its arguments (intuitively, the ‘core’ davidsonian event) *together with* all the adverbial, prepositional and modificatory devices at a language’s disposal.

¹ For example, they contain at most one force-dynamical change, at most one direct cause, and at most one specified result state. See Ramchand (2008) for discussion.

Now, while full propositional content at the sentence level can indeed be modeled by situations (as argued initially by the pioneering work in Barwise and Perry 1983), it is a separate step to say that situations should be given status as part of our object language in semantic description. The second step does appear to me to be warranted, and the evidence for the ontological reality of situations is persuasive. As noted in Kratzer (2014) ‘examples can be constructed to show that natural languages have the full expressive power of object language quantification over situations’. Situations and Austinian ‘topic situations’ (Austin 1950) seem to be needed to account for:

- (i) truth conditions in context,
- (ii) tense marking (Klein 1994), and
- (iii) are necessary in a variety of ways for quantifier domain restriction (see Kratzer 2014 for details).

I will take it then that we have linguistic evidence for the reality of event descriptions from the core properties of verbs and verbal meanings, and we also have evidence from a wide variety of discourse level effects for the reality of situations. But are Davidsonian events and situations the same thing? And what is the relationship between the verb denotation and the rich situational description that eventually gets established at the sentence level?

The answer given in Kratzer (2014) is that events and situations are indeed exactly the same ontological type, but that events are minimal situations. In her view, the notion of exemplification mediates the relation between propositions and Davidsonian events, and makes explicit how the latter relate to situations more generally. I repeat the definition from Kratzer (2014) below in (1).

(1) **Exemplification**

A situation s exemplifies a proposition p iff whenever there is a part of s in which p is not true, then s is a minimal situation in which p is true.

The intuition is that a situation is something that propositions can be ‘true in’, but a situation *exemplifies* a proposition if it is the minimal such situation, with no extraneous, unnecessary parts. It is the ‘minimal’ situation that makes the proposition true.

If we incorporate Davidsonian event semantics into situation talk in this way, we get (2) as the representation of a sentence such as *Ewan swam for 10 hours.*, which is taken from Kratzer (2014).

$$(2) \quad \lambda s[\text{past}(s) \wedge \exists e[e \leq_p s \wedge \text{swim}(\text{Ewan})(e) \wedge \text{fhour}(e) = 10]]$$

So the sentence here is a property of situations such that the situation is ‘in the past’ and there is an event that is a subpart of it which is the exemplification of

Ewan swimming. The temporal measure of the exemplified event in hours is 10. Here we see that both situations and events can be arguments of temporal modifiers. One could also break down events/situations into their temporal run times and explicitly predicate the temporal predicates of these intervals instead, *or* we state it as above and allow the specification of the temporal predicate to make this precise. The important point here is that events/situations in this system have temporal parameters and temporal properties can be ascribed to them.

According to Kratzer (2014), the formula above ".... incorporates the usual notation for Davidsonian event predication. Within a situation semantics, this notation is just a convenient way to convey that *swim(Ewan)(e)* is to be interpreted in terms of exemplification: we are not talking about situations in which Ewan swims, but about situations that *exemplify* the proposition 'Ewan swims'." (my italics).

For Kratzer, an event is ontologically the same kind of animal as a situation, but it is one that stands in the exemplification relation to a particular kind of atomic proposition (namely, the ones that we usually assume are the introducers of Davidsonian events). So in this way, Kratzer is in fact relating the use of the term event to something independent *about the syntax*. In effect, if a proposition comes from the interpretation of vP then it corresponds to 'event', while if it comes from a larger syntactic phrase then it corresponds to 'situation'. This however has no real effect on the *semantic* ontology.

The problem with the standard view as described above is that it underplays the differences between the semantics of the inner vP (what I have elsewhere called the 'first phase' Ramchand 2008) and the higher parts of the clause. To understand what I mean by this, it is necessary to briefly discuss certain typological patterns in linguistic forms.

1.1 Linguistic Generalizations and Constraints on the Syn-Sem Mapping

It is a truism perhaps that the syntactic representations of natural language need to be given a compositional interpretation (cf. Heim and Kratzer 1998). However, the standard mechanisms used in formal semantics for modeling compositional interpretation are in fact extremely powerful. The unfettered lambda calculus, endowed with abstraction over predicates of higher types can put any jumble of words or structures together to deliver the final desired output reflecting our description of the intuitive truth conditions (see Higginbotham 2007). The notation itself overgenerates in the absence of explicit constraining principles. Explicit constraining principles after all are the job of linguists, not of the notation.

So what constraining principles do we need? Ideally, in my view, we need to build a compositional semantics of the clause that makes the deep and uncontroversial generalizations about verb meaning fall out as a natural consequence. In Ramchand (2016), I argued that we need the compositional semantics of the vP to reflect the universal hierarchical structuring of causal embedding. In this monograph, I pursue the logic further and take seriously the typological fact that natural languages universally encode temporal information hierarchically *outside* of the causal and force dynamical content of the event itself. This universal fact about semantics is rarely perceived as such because it has already been reified as ‘syntactic’ fact in the form of a phrase structure template: CP > TP > VP. The template with these three zones is as much a template as any more articulated cartography (cf. also Ramchand and Svenonius 2014 for discussion), and at this stage of our understanding simply has to be stipulated. It follows from nothing else. Unfortunately, it also does not fall out from event semantics, under any current understanding of the term. This is because, on current understanding, events (and situations) as well as trafficking in notions such as ‘causation’ and ‘agent’, also have properties related to time because they are particulars with a particular time course.

Consider a hypothetical language spoken on the planet Zog. The planet Zog is a world very different from our own, inhabited by many strange living creatures, one species of which have acquired symbolic thought and speak their own form of language. This is Zoggian, and it has properties found in no human language. In particular, we find the bound morpheme *fub*² which denotes roughly ‘the process of dissolving into a green slimy puddle’. In addition, we find the bound morpheme *-ax* which has the semantics of PAST and the bound morpheme *ilka* which has the semantics of CAUSE. Like human languages, Zoggian works by generating hierarchical symbolic structures with predictable interpretations. However, unlike Human, the PAST morpheme always occurs hierarchically closer to the conceptually rich part of the verbal meaning than the CAUSE morpheme does. The relevant sentences of Zoggian follow in (3). (Note also in passing that Zoggian’s basic word order is OSV).

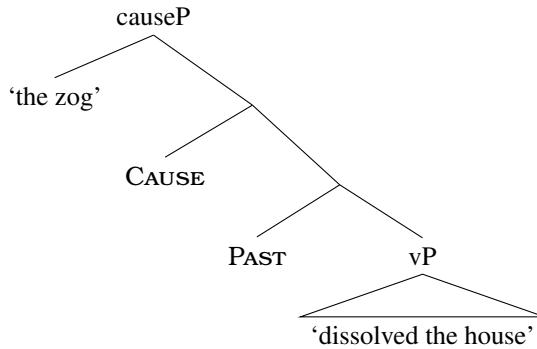
(3) blix a fub-ax
 the.house dissolvegreen-PAST
 ‘The house dissolved into a green slimy puddle.’

² As a simplification, and for the purposes of exposition, I will translate all Zoggian forms into IPA, and use descriptive terms from human linguistics. In fact, Zoggian does utilize an auditory channel but one which is not perceptible to the human ear.

(4) blixā marrg fub-ax-ilka
 the.house the.zog dissolvegreen-PAST-CAUSE
 The zog dissolved the house into a green slimy puddle.

The tree structure for the sentence in (4-b) is given below in (5)

(5)



Suppose further, that it turns out that there are many Zoggian language families but with very few exceptions, Cause appears external to temporal information. This is no problem for a compositional semantics. Indeed, it is no problem for the semantics developed for Human languages either. For example, simple denotations for the Verb, Cause and Past morphemes could be given as in (6) below.

(6) (i) $[[vP]] = \lambda e[fub(e) \wedge \text{Undergoer}(e) = \text{'the house'}]$
 (ii) $[[\text{PAST}]] = \lambda e[\tau(e) <_t \text{'now'}]$ (where τ is e 's temporal trace function)
 (iii) $[[\text{CAUSE}]] = \lambda x \lambda e[\text{Causer}(e) = x]$

The vP combines with the PAST morpheme by argument identification to give:
 $\lambda e[fub(e) \wedge \text{Undergoer}(e) = \text{'the house'} \wedge \tau(e) <_t \text{'now'}]$

This then combines with the cause morpheme, again by argument identification to give:

$\lambda x \lambda e[fub(e) \wedge \text{Undergoer}(e) = \text{'the house'} \wedge \tau(e) <_t \text{'now'} \wedge \text{Causer}(e) = x]$

Note that the denotation I have given for Cause is essentially identical to the compositional system assumed by Kratzer (1996) for the folding in of the external argument by argument identification. The only difference with Zoggian, is that I have folded in the meaning of past tense first, also by argument identification.

To return to English, the analogous case would be verbs that looked like *slimed/slimedify*, instead of the more natural *slimify/slimyfied*.

(7) (Zoggified English)/Zoggedify English

- (a) The house slimed
- (b) The man slimedify the house.

It is important to understand why the semantics works. The trick is not the use of argument identification itself (what Heim and Kratzer 1998 call ‘predicate modification’). As Heim and Kratzer (1998) demonstrate, argument identification can always be rewritten as some form of function composition, although the formulas look less transparent. The reason for the commutability of *PAST* and *CAUSE* is that the new information is added by a simple conjunction. Argument identification is not sensitive to what it is being added to, as long as it is also a legitimate predicate over events. This means that regardless of the order in which you combine *CAUSE* and *PAST*, you end up with the same representation. And regardless of how you tweak the subtle details of the *CAUSE* and *PAST* denotations, this will always be the case just as long as *both of those predicates manipulate semantic factors that are properties of events*. If events are temporal entities, then *CAUSE* and *PAST* both denote properties of some aspect of the event. The only way to enforce one ordering over another would be to stipulate a presupposition for *CAUSE*, for example, that it can only combine with event descriptions that have not yet been located in time. While lexically specific information like this is certainly attested in natural language, it seems like we want to have a more general way to approach a human language universal.

Semantic theory can do a little bit better if it imposes a different structural type on tense predicates. We could stipulate that even though events have temporal properties, the semantic type of functional tense is always such that it relates constituents that are properties of times directly, and that this switch comes after the lower *vP* domain. This would be a way of ensuring that grammaticalized tense formally composes semantically *after* cause, as indeed in much recent semantic work (Kratzer 2000, Klein 1994). The point I am making here is only that there is nothing about the internal logic of events that makes this stipulation natural or obvious. It can be made, but it is the equivalent of the syntactic stipulation, achieving the ordering by fiat. We could *imagine* things to be otherwise, but they never are. The current ontology which utilizes events which are full blown temporal (and even worldly) particulars does not naturally underwrite the ordering facts we find in the *syntaxes* of human languages. Something needs to be stipulated on top of this system in any case, and the question is exactly what. Basically, I am not satisfied with a

templatic approach to the problem as delivered by the functional sequence, or by the semantic correlates in terms of stipulated distinct types. The functional sequence (even in the form of the very pared down CP > IP > VP template) is a convenient descriptive device that should be anchored semantically in some deeper explanation. Stipulated semantic orders also seem to me to not go deep enough, and are also essentially templatic. I would like to use the robust empirical facts about ordering to motivate a restructuring in the semantic system of composition from the inside, one which actually does have the well known templatic facts about sentence semantics and hierarchical ordering as a consequence. I am looking for an explanation of the templatic effects from the semantic and cognitive primitives that form the basis of natural language. The hope is that by being explicit about how the semantics needs to be structured, we make a start on filling in one side of the equation for the communication with C-I system, the more general cognitive mechanisms of mind/brain.

Note also that I will be essentially reversing the normal methodology found in much work on the syntax-semantics interface, which starts with establishing the explicit and precise representation that has the correct truth conditions (since this is something that is accessible to verification by intuitions), and *then* partitions those contributions among the elements of the sentence in some kind of function-argument notation. My methodology will be to start with the individual building blocks and come up with a reasonable description of their contribution to the sentence meaning. I will then try to build up gradually to a final semantic representation that has the right truth conditions. The two strategies should in theory give rise to the same syn-sem analysis, but in practice they do not. This is because the one strategy (i.e. my own) places more weight on the integrity of different lexical items and their constancy across instances of use, and the other (the standard semantic methodology) places more weight on expressing the full precision required for determinate truth conditions.³

As part of this agenda, I take seriously the robust crosslinguistic generalization that tense and aspect inflection when they appear (and they seem to in over 80 percent of the worlds languages)⁴ are hierarchically *outside* of the core

³ This is absolutely not to disparage representations with clear and precise truth conditions, but rather to take into account the fact that syntactic representations simpliciter may in fact underdetermine those more explicit representations, depending on the division of labour between grammatical ingredients and contextual and pragmatic effects. The methodology I will be adopting is conservative in what it ascribes to the grammatical system.

⁴ In WALS (Dryer and Haspelmath), it is reported for their database of languages that only 31 out of 201 typologically distributed languages had no past, future or aspectual inflection on the verb. (Whether a language has tense marking or aspect marking turns out to be independent, contrary to folk belief). See also Julien (2002) for a typological description of the relative ordering of voice and cause morphology with respect to tense and aspect.

verbal description (including the description of cause, process and result in the verb). Note that the existence of a template in this sense is not falsified by languages that do not have the appropriate inflection (since there is no claim of universal overtness here), but by a language that has morphemes with the relevant semantics but in the wrong order. So although one *could* imagine the reverse (and we would be able to build a compositional semantics to describe it as we saw for Zoggian), this does not actually seem to happen in the worlds' languages. In other words, verbs do not share stems for tense specification, with suffixes and/or prefixes indicating the specifics of the description for the dynamic process or stative situation involved.

Robust generalizations of this type, at such a basic level of language structuring, are impressive and need to be taken seriously. The use of situations to model propositions does not at present give us any purchase on why the template should be the way it is. However, situations are an important step forward because they have the right structure to subsume the kind of information and modifications that we find in the verbal extended projection, namely reference to times, worlds, locations, and ultimately the speaker.

1.2 Event Properties and Event Instantiations

To summarize, the Kratzerian notion of situation has the following definitional characteristics (See Kratzer 1989 for detailed exposition).

- Situations are particulars that exist in worlds and at times
- Situations stand in part-whole relationships to each other (they form a mereology). In other words, situations can be temporal or spatial subparts of other situations. (We will use the notation \leq_s to express situational subpart).
- Situations grow deterministically into a particular possible world, in the sense that there is a unique maximal element in the mereology. So situations are just instantiated world parts, and the maximal element in a situational mereology is a particular world.

So every s is related to a unique possible world w . We can say then that situations have world parameters, since they are deterministically related to a particular world. They also have time properties, or parameters, in this sense, and they can also be ‘proposition-exemplifiers’, and more complex situations can be related to simpler Davidsonian atomic exemplifying situations via the subpart relation. I will largely stick to this view of situations as worldly particulars located in time, and I will also have use for the intuitive correspondent to Kratzer’s notion of ‘exemplifying situation’ at the level of the vP.

Within the vP, however, I will argue that we need a rather different sort of beast than (even a minimal) situational description. Intuitively, we will need

representations that express force-dynamically relevant descriptive content and relationships to participants, but which do not have temporal (or world) information. But how can we even formalize what it means to be an event without making reference to being in the world, and therefore being part of a particular world and time?

1.2.1 Philosophical Antecedents

There is a long tradition of formal semantic work which grounds all of our semantic definitions in objective particulars. This work leads up through Quine and Carnap and culminates in Lewis, perhaps the most influential of the analytic philosophers on modern semantic thinking. In this kind of tradition, atemporal and awordly statements are built from the mundanely worldly ones, by generalizing over times and ‘possible’ worlds. This hugely successful and productive intellectual tradition lies at the heart of our classical semantics of propositions and modality,

However, there is an alternative strand of thought that is troubled by the feeling that this is often simply not an adequate rendering of our intuitions about linguistic meaning. For example, the philosopher Kit Fine argues that we need to make a distinction between facts that are always true as a matter of contingent circumstance, and facts that are *essentially* true.

“Empiricists have always been suspicious of modal notions. For them, the world is an on-or-off matter—either something happens or it does not; and there appears to be no room in their on-or-off world for a distinction between what happens of necessity and what only happens contingently or between the essential features of an object and those that are only accidental . . .

Fine (2005)

Fine here is interested in the fact that certain inherent or essential properties of an object are there in a logically prior way, and not as a ‘contingent’ fact about how the world is. We feel intuitively there is a difference, but in the end the possible worlds way of expressing the idea of essential properties in terms of ‘*all* possible worlds’ gives technically the right results in many cases, but also seems to obscure the intuitive difference that we feel between the essential truth and the contingently universal truth. He claims that semanticists in this case have been willing to sacrifice this intuition in the service of an idea that is more tractable formally and mechanically. He continues.

For empiricists, in so far as they have been able to make sense of modality, have tended to see it as a form of regularity; for something to hold of necessity is for it always to hold, and for something to hold possibly is for it sometimes to hold. But if there is not enough going on in the actual world to sustain the possibilities that we take there to be, then one strategy for the empiricist is to extend the arena upon which the possibilities are realized to include what goes on in each possible world. Of

course, such a view is compatible with a moderate realism in which possible worlds, and what goes on in them, are taken to have a different ontological status from the actual world and what goes on in it. But combine the regularity view of modality with a nominalism about what there is and we end up with a position very like Lewis's. Indeed, it might be argued that, *au fond*, Lewis is as sceptical of modal notions as Quine. Neither can understand modality except as a form of regularity; and the only difference between them lies in the range of the regularities to which their respective ontologies allow them to appeal....

As we will see more directly when we get to the chapter on modality, the correspondence between quantification and modalized propositional meaning is very seductive. There is clearly something deep at stake here and logical quantification is one way of getting at the mystery of hypothetical reasoning. However, it might not be the only way. In the case of 'essential truths', we might decide to begin with these as primitives instead of recasting them in extensional notations. Fine in particular argues that transcendental essences in his sense have to exist in a *basic* form, in a separate dimension of existence not derived from or defined in terms of worldly instantiations (both real and possible). I quote from 'Necessity and Non-Existence' here.

"Finally, it will be suggested that the identity of an object— what it is— is not, at bottom, a worldly matter; essence will precede existence in the sense that the identity of an object may be fixed by its unworldly features even before any question of its existence or other worldly features is considered."

I will argue that what we need here in the representation of events is similar in that *essence* must precede *existence* in the cumulative building up of a natural language proposition. Fine's own arguments concern objects and identity and he explicitly does not carry those arguments over to the discussion of events as objects. Empirically, I think the arguments can be made for the event domain based on phenomena like the imperfective paradox and resultative participles. I also think that the linguistic evidence from verbal templates complements the philosophical arguments from Fine. Under plausible assumptions about the mapping between syntax and semantics, the cumulative hierarchical complexity of syntactic structuring should be paralleled by increasing semantic complexity. I think it suggests that semantically also, at least for the human mind, the verbal concept and its arguments is conceptually prior to its embedding in time and world. Linguistic facts paint a consistent picture across human populations to this effect, it would be nice if the semantic composition reflected this consistent picture.

To put the argument another way, the grounding of *all* denotational types in worldly meanings (by which I mean particulars in worlds and times) by using worldly events/situations at every level of the clause, would make it impossible

to derive the layering of syntactic structure by purely semantic means. Specifically, the fact that time reference provides the *outer clothing* of verbal meaning in language after language and never the other way around, is something that would have to be stipulated as a universal syntactic fact. If we wish to reduce syntactic stipulation and see explanations for deep typological generalizations in facts about cognition, then we need to adopt a semantic framework that is more sensitive to the patterns that syntax gives us. Only then will we have the vocabulary which is commensurate with the primes of psycholinguistic and neurolinguistic investigation, and with which to address the interface questions. At the very least, I think it is worthwhile approaching the problems of meaning and the interface with cognition from the inside out, by looking at the ways in which language structures the way meanings are built up.

This also returns us to an old debate concerning internalist vs. externalist theories of meaning. The debate has already been won by the externalist camp on the entirely reasonable grounds that there is a crucial 'aboutness' to language, and that if we attempt to ground our theories in internalist notions then we are condemned to theories that make no sense of the inter-useability of language and which end up being at best unfalsifiable and mystical at worst. For this reason, our theories are grounded in notions of reference to individuals in the world, and to truth. However, I find myself in agreement with Chomsky (1995) where he convincingly shows that reference and truth are themselves mystical notions, and that the quest to fill out some materialist agenda closing the gap between mind and body in itself presupposes an unwarranted dualism. Citing Nagel 1993, Chomsky argues that:

‘ It is a hopeless task to "complete the materialist world picture" by translating accounts of "mental phenomena" in terms of a "description that is either explicitly physical or uses only terms that can apply to what is entirely physical" or perhaps give "assertability conditions" on "externally observable grounds". (Nagel 1993. pg.37) Chomsky 1995. pg 4 ’

Instead, what he advocates is a naturalistic internalist view. Naturalistic in the sense that we attempt to study humans just like anything else in the natural world, but internalistic because we are attempting to understand the internal states of an organism. This is not to say that we could not and should not have a methodologically naturalistic approach to what we think of as 'external' and 'objective'. Here is Chomsky (1995) again.

“ One branch of naturalistic inquiry studies common sense understanding. Here we are concerned with how people interpret object constancy, the nature and causes of motion, thought and action, and so on ("folk science" in one of the senses of the term). Perhaps the right way to describe this is in terms of beliefs about the constituents of the world (call them "entities") and their organization, interaction and origins. ... It is an open question whether, and if so how, the conceptual resources of folk science relate to those involved in the reflective and self-conscious inquiry found in every known culture ("early science") and to the particular enterprise we call "natural science". For convenience let us refer to the study of all such matters as

"ethnoscience". It is also an open question how the conceptual resources that enter into these cognitive systems relate to the semantic (including lexical) resources of the language faculty. . . . The ethnoscienctist seeks to determine what people take to be constituents of the world, however they may talk about it. A different inquiry seeks the best theory of language and its use, and the states, processes and structures that enter into it.

Chomsky (1995), pg 28-30.

I consider this book to be part of the second line of inquiry described by Chomsky above. It takes seriously the idea that natural language symbols correspond to mental representations (formed and codified through experience of the 'world'), which are then deployed by human beings in a particular context to help describe other particular things they take to exist in the world.

"Neurologist Rodolfo Llinás (1987) puts the matter well when describes perception as "a dream modulated by sensory input", the mind being a "computational state of the brain generated by the interaction between the external world and an internal set of reference frames". But the internal frames that shape the dreams are far more intricate and intriguing than often assumed, even at the level of the lexicon, still more so when we turn to expressions formed by the computational processes."

Chomsky (1995), pg 23

In Ramchand and Svenonius (2014), we argued that a new ontology of sorts was required, one that contrasted atemporal and aworldly 'event essences' with 'situational' particulars. The new member of the ontology, intuitively the equivalent of essential conceptual events, or 'event types' is a slippery notion which is difficult to make coherent without resorting to possible worlds themselves for their definition. A new ontological primitive for language and its integration into a compositional semantics that still grounds itself in the external significance of language that we attempted to motivate in Ramchand and Svenonius (2014) poses problems of a foundational complexity that are hard to solve. Fine (2000) takes a different approach and presents a formal semantics for his logic of essence which uses a special notion of truth with respect to a predicate F , \Box_F which is truth *in virtue of the nature of that predicate*.

It is also relevant to point out some current semantic research that is pursuing this kind of intuition in terms of event *kinds*. I refer here to recent work by Louise McNally and Berit Gehrke who argue based on certain kinds of nominalized and participial forms that the event kind actually forms the *basis* of all subsequent event denoting forms (Gehrke and McNally 2015, Gehrke 2015, Grimm and McNally 2015). I consider this to be convergent research, albeit employing a more standard formal system. One difference with that work comes from the fact that I will claim a 'concept'-like notion operates in principle throughout the first phase. Another difference is that, instead of invoking a primitive corresponding to event kinds, I will be using an implementation in

terms of ‘partial event descriptions’ and link the nature of that partiality to the nature of the conceptual content of lexical as opposed to functional items, a move that will end up having its own technical consequences.

The solution to reconciling the intuition that natural language symbols denote essential mentalistic concepts (Fine, Chomsky) with the necessity of linking up with the ethnoscientific constituents of the ‘world’ to convey aboutness and interdescribability will require a drastic change in the system. We must reify the contextual and self conscious aspect of meaning and build it into the system of representation itself. I will propose an implementation of that intuition by using *words of the language as elements in the domain of individuals*, inspired by Potts (2007) who first proposed it for the analysis of metalinguistic comment, and used explicitly in the analysis of ideophonic elements in Henderson (2015).

The link between words of a language and their semantics is achieved in acquisition mediated by perception/cognitive uptake of events in the external world, but they are, crucially, generalizations across particulars which can then be deployed by a speaker. In this way, the lexical item is the codification of a certain implicit perceptual and cognitive generalization, *reusable* as a bridge between internal representations and external events.

It is inspiring also in this regard to recall the work of Barwise and Perry (1983) on situations, who cite the efficiency of language as one of its central design features. In other words, alongside the external significance of language and its productivity and compositionality, we also have the fact that words of a language are efficient members of the code in being able to be *reused in situation after situation*. To this they add correctly the perspectival relativity of language, its ambiguity, and its mental significance.

So, for human language to get off the ground, we need to be in possession of symbols that are *shape abstractions over the different actual situations encountered in the learning phase*, and a speaker is then able to deploy those symbols as a means of characterizing new situations in the world as she comes across them. The reusable highly efficient part of semantics is what Barwise and Perry call ‘meaning’, and at this level the symbols are compositional and efficient, but they greatly underdetermine the information conveyed by a sentence. To get from abstract and reusable meanings to actual information, we need to know facts about the user, the *deployer* of the meanings and where they are located in time and space. In turn, the information conveyed by a sentence is related to the actual situation by the description relation— a situation can be (partially) described by some information, or conversely, information can be supported by a situation.

In terms of implementation, my inspiration has come from the apparently extreme and exotic case of ‘ideophones’. Henderson (2015) states that work on the formal semantics of ideophones is scarce because of the ‘difficulty in formalizing the distinction between descriptive meaning and depictive meaning, which ideophones seem to traffic in’. In giving his own account, Henderson explores a formal foundation for the notion of *demonstrations* from Davidson (2015) and extends it to account for the ideophonic data. According to Henderson, demonstrations are a special type of communicative event that stand in a similarity relation with the event demonstrated. It might seem like a curious and typologically rare corner of the semantics of human language, but in fact, I think that the problem posed by ideophones is perfectly general, it is just that it can be seen most clearly in this extreme case. In the general case, we need to link the properties of the communication/demonstration event with the symbols being actively deployed in order to achieve the description of a real world particular.

The idea is the following: the lexical predicates of a language are reusable symbols of event depiction encoding a cognitive/perceptual semantics, derived from experience of the world but involving primitive and natural generalizations over actual physical instantiations. Because these depictive semantic symbols are formed by learning and labelling *generalizations* over time and space, they are by definition silent about temporal and locational properties. They involve the recognition of specific causal and force dynamical relationships among actants (see also Copley and Harley 2015), and the ability to ascribe basic cognitively apprifiable properties to individuals.

In an actual proposition, a speaker *uses* these depictive predicates to abstractly characterize an actual existing event, and then only subsequently adds information about that event’s temporal and locational properties. These latter properties are defined in relation to the speaker’s own *Origo*.

In short, the standard classical semantics involving event particulars does not do justice to the order in which natural languages build up propositions from linguistic symbols. In the semantics proposed in this monograph, lexical items need to be taken seriously as individuals deployed by a speaker. These lexical items are bundles of form and meaning which contain cognitive abstractions of event properties that transcend time and place. The choice of the primitives of these abstractions will be motivated by the data on verb meaning and classification. This way of doing things does not immediately improve on certain open questions concerning conceptualization, but it does offer a system which will be able to connect more systematically with the syntax on the one hand, and with the units of cognition and language processing on the other.

1.2.2 A Quotational Semantics for Natural Language

Recall that we are going to build up a representation of the proposition in three stages:

- (i) The putting together of lexical items which encode certain event properties. This stage needs to be productive and compositional, but with no reference to temporal or world parameters.
- (ii) The assertion by the speaker of the existence of an event in world and time with those properties.
- (iii) Addition of specific temporal and world properties to the event.
- (iv) Anchoring of the worldly and temporal properties via the *Origo* (the speaker and her contextual coordinates).

In order to do this we need to add to the usual model, a domain D_μ which is the domain of well-formed linguistic entities of type μ , after Potts (2007). These linguistic objects are triples, consisting of a \langle phonological string, syntactic features, SEMANTICS \rangle . Full expressions of type μ will be written in sans serif .

So for example, the verb *run* might have the denotation:

$$[[\text{ run }]] = \langle \text{ run}, \langle \text{init, proc} \rangle, \lambda e[\text{run}(e)] \rangle$$

For convenience, we adopt the convention in Henderson (2015) which uses the bottom corner notation to pick out the semantic part of the triple denoted by something of type μ .

$$\text{Thus, } \llcorner \text{ run } \lrcorner = \lambda e[\text{run}(e)]$$

We can think of the building up of a proposition as a deployment of something of type μ , to create a relationship between a (complex) event property and a demonstration event.

In Ramchand (2008), I argued that the compositional relationships among lexical items in this domain are restricted to certain basic relations of causation and result, and also to the **HOLDS** relation that relates subevents to their actants. The lexical conceptual content of verbs obviously varies without limit in other dimensions. Crucially however, the lexical conceptual content of verbs cannot and does not include temporal information, since these are cognitive *abstractions over time and space*. Apart from this part of the claim, it is not the purpose of this monograph to investigate or argue for a particular view of the internal semantics of event descriptions. The way the system is set up here, it in principle allows the reader to slot in their own denotations for lexical items and the functional projections within the first phase. The important part of the system for our purposes is that we need to have methods for composing elements of type μ , to create derived elements of type μ by the end of the first phase.

The rule for composing μ entities will be the following:

Box 1.1
Language Symbols as Objects in the Ontology

(8) (a) Symbols of the language constitute the domain D_μ , which are triples consisting of a \langle phonological string, syntactic features, semantics \rangle
 (b) The semantics of a verbal LI are partial descriptions based on sensory and cognitive abstractions over experience.
 (c) The syntactic part of the information in a triple that is a member of D_μ , is a subtree of the language. The merge of $u_1 \in D_\mu$ and $u_2 \in D_\mu$, creates a derived element of D_μ , u_3 , which has the syntactic representation built by merging the syn-rep of u_1 with the syn-rep of u_2 , and a semantics is composed by ordinary argument identification of $\llcorner u_1 \lrcorner$ and $\llcorner u_1 \lrcorner$.

The linguistic unit so formed will also be an ordered triple and its phonology will be formed by concatenating in some way the phonologies of the two inputs⁵. The syntactic representation will also presumably be composed via some algorithm, but once again this is not directly the issue for us here. The important thing is how the semantic parts of the triple compose. In the simple cases I begin with, the semantics parts of the triple will compose in the normal way by argument identification (simple conjunction of properties).

After the completion of the lexical symbolic part of the syntax, I will assume that a functional item, which I will call Evt is merged with the result. This is my name for that head that introduces the generalized deployment operator, and is at the edge of this first phase of building. The Evt head deploys the lexical content built up so far and creates something that now denotes a property of events directly via the introduction of the demonstration event d. I take the use of d from the formalization in Henderson (2015), performing act of communication. It is similar to the Kaplanian context c, but conceived as a ‘Davidsonian’ situational variable corresponding to the utterance event.⁶

⁵ This monograph will have nothing to say about linearization

⁶ The use of the Davidsonian event variable corresponding to a performative verb in Eckhardt (2012) also bears a close relationship to this idea. For Eckhardt, this variable (which she labels ε) denotes the ongoing act of information transfer. The existence of this variable is necessary for the analysis of performatives and adverbials like *hereby*, although not sufficient, since the analysis of performatives also requires an explicit definition by the speaker of what her utterance is doing. In my own system here, d is represented explicitly as the ongoing act of information transfer, and does not require an explicit performative verb for its introduction. One could see the quotational system as a hybrid version of an analysis where every utterance is preceded by an implicit performative—‘John is tall.’ is really ‘I ASSERT THAT John is tall’, except that no literal embedding or deletion

This is Henderson (2015)'s denotation for the quotation meaning. $\text{TH}(d) = u$ says that the 'theme' of d is the linguistic object u , and d 'demonstrates' or has certain structural properties in common with e .

(9) QUOTE : $\lambda u \lambda d \lambda e [\text{TH}_\delta(d) = u \wedge \text{DEMO}(d, e)]$

The central idea of my own adaptation is that the notion of demonstration is simply a special case of the more general idea that the speech event d is used to CONVEY an event in the world e . The deployment of a lexical item as the thematic content of d is also perfectly general. This very general schema underlies both acts of description and depiction and indeed everything in between. Thus, more generally we want to say that symbolic content is deployed by the speaker to convey an event. For concreteness, I define the Evt head at the edge of the first phase as introducing the utterance situation d , with linguistic content u in order to convey event e .

Box 1.2
Deployment of the Symbolic Content at EvtP

(10) I. EvtP : $\lambda d \lambda e [\text{UTTERANCE}(d) \wedge \text{TH}_\delta(d) = u \wedge \text{CONVEY}(d, e)]$
 Property of an utterance event d and event e , which has u as its theme, and where d is deployed to convey e (where $u \in D_\mu$ is the denotation of the first phase verbal description).
 II. In the case of purely conventional (i.e. non-depictive) LIs, uttered with sincerity and without metaphor or hyperbole,
 $'\text{TH}_\delta(d) = u \wedge \text{CONVEY}(d, e)' \longrightarrow ' \sqsubset u \sqcup (e)'$

In the specific case of quotes and iconic items and ideophones the CONVEY relation becomes Henderson's (2015) DEMO relation. More generally, though, in the case of straightforward deployment of a linguistic item with no imitation or iconic elements in d , CONVEY simply reduces to the event e that d is demonstrating having the semantic property encoded by u . Thus, in the case of the EvtP built at the edge of the first phase in English, we get the following denotation.

(11) EvtP : $\lambda d \lambda e [\text{UTTERANCE}(d) \wedge \sqsubset u \sqcup (e)]$

is taking place. Rather, the claim will be that a representation of the eventuality corresponding to the ongoing utterance event is explicit in the compositional semantics, and is necessary to convert the usage of symbols of D_μ into something that has explicit truth conditions.

To return to the motivations and justification of this move, we note that it is a representational encoding of the intuition that reference involves a speaker and a context in addition to the symbol she is deploying. But it is not just a matter of a speaker X using the symbol Y to refer to the object Z, we need to leave room also for the contextual circumstances and mode of deployment of the symbol in question. Once again Chomsky (1995) puts it more accurately,

"More generally, person X uses expression E with its intrinsic semantic properties to talk about the world from certain intricate perspectives, focusing attention on specific aspects of its, under circumstances C, with the "locality of content" they induce (in Bilgrami's sense). "

Chomsky (1995), p. 43

Within the external clothing of speaker deployment however, we can still maintain quite a conservative system of semantic composition, with verbal elements in the first phase simply denoting certain event properties. An important feature however is the stipulation that all event properties encoded in lexical items are *generalized abstractions and do not have any temporal, worldly or locational properties*. They are thus partial descriptions that reflect the idea of 'essential' properties or 'event concepts'. Technically speaking, in the lowest domain, the semantics just composes elements of type μ , in ways that we will see more of in chapters 2 and 3. It is at the Evt level that we introduce the generalized equivalent of a quotation operator, and crucially, in doing so, introduce the enclosing demonstration event, which I take to be Davidsonian event corresponding to the utterance.

At this point, it is worth pointing out again the differences between this system and the ideas being pursued in the related work of McNally and Gehrke (Gehrke and McNally 2015, Gehrke 2015, Grimm and McNally 2015). The problem with introducing event kinds as primitives is the work that needs to be done to state the conditions on composition at this level. For example, in Gehrke (2015), kinds and their subkinds are related by prototypicality relationships. There is an intuition there seems essentially correct but it is hard to pin down, especially since it seems to me that event kinds can be extremely internally complex and indeed, novel. In the treatment I am proposing here, the illusion of lexical genericity in this sense is a byproduct of the fact that symbolic members of D_μ are partial event properties that are *abstractions over space and time*. This means that the LI itself invokes only those properties that are independent of instantiation, by definition. The system then allows event properties to be added to and composed via a normal kind of Davidsonian conjunctivism.

Once the quotation operator has been introduced, we are in a position to define temporal and world operators. Recall that the central problematic we

face in capturing our typological generalizations is that certain types of information robustly occur outside others in the syntactic representation. We have constructed a domain of composition of lexical items which composes properties of events. What stops us from having temporal properties of those events in the lexical items denotations? The answer I will give is that temporal and locative predication are not properties of events, but *relations* between events and deployment events d . This means that temporal information is simply not statable until Evt is merged at the edge of the first phase.

Further, I will follow Champollion (2015) in introducing the closure of the event variable low down, with the merge of what I will call the *Asp* head. I choose the label *Asp* here for the point of closure of the event variable, to emphasise the convergence with work on the syntax-semantics of tense and aspect, namely that this is the position where the switch from events to times happens. In other words, empirically it seems as if the transitional point occurs precisely here at the left edge of *vP*. I think the observation and intuition from previous work is essentially correct, but I will embed that intuition within a quotational semantics and incorporate Champollion's proposal about event closure and event properties.

Champollion (2015)'s reason for introducing what he calls 'quantificational event semantics' is that traditional event semantics sits rather uneasily beside certain other welcome results and generalizations in formal compositional semantics. This has led some semanticists to reject formalisms using the event variable and try to rework the system in other terms (see Beaver and Condoravdi 2007 for a recent proposal). The problem with events arises because the event variable, although treated like any other object variable for some purposes, has a curious relationship to other quantified variables within the sentences. Specifically, the event quantifier itself never interacts with any other quantifiers—it always takes narrow scope with respect to them. The innovation that Champollion proposes is to take verbs themselves to denote sets of sets of events. Essentially, verbs and their projections denote existential quantifiers over events, and the event variable is no longer considered to be bound at the sentence level as in standard accounts. Once this move is made the rest of the semantics can be business as usual. Here is Champollion's denotation for the verb phrase *see Mary*.

$$(12) \quad [[\text{see Mary}]] = \lambda f \exists e [\text{see}(e) \wedge f(e) \wedge \text{th}(e) = \text{Mary}]$$

The verb phrase now denotes a property of event properties, a move that is required to allow further properties of the event to be added after existential closure. Champollion dubs his approach "quantificational event semantics" and I will follow his insight here. The quotational approach is a species of

quantificational event semantics in which the existential binding of the event variable is mediated by the introduction of the demonstration event at the edge of the first phase. I will label the locus of existential binding for e , Asp . As in Champollion's system, the semantic type of AspP is a property of event properties (his ' f ' in (12)). In the quotational implementation I will use f_d for my variable over event properties to indicate that these properties are essentially relations between an event and an anchoring utterance event d .⁷

In more standard theories, the *Origo* is fundamental part of the model that interprets semantic representations. The quotational theory reifies this as a part of the representation. Correspondingly, the notion of temporal and world information are now explicitly world-temporal *relations* between d and the e that is being described. I will use f as this variable over spatiotemporal relations between e and d . Intuitively f is the class of relations that locate an event to an utterance situation, where location can be spatial, worldly, or temporal—the core indexical parameters of that situation. We are now in a position to express the meaning of the Aspectual head that the EvtP combines with. Asp looks for a property of demonstration events for an event e which existentially binds that event variable and creates a property of spatiotemporal properties of e (rooted in d). The AspP built up by the quotational quantificational system, will therefore look as in (13).

$$(13) \quad [[\text{AspP}]] = \lambda f_{<v, <v,t>>} \lambda d \exists e [\text{Utterance}(d) \wedge \sqcup u \sqcup(e) \wedge f(d)(e)]$$

⁷ Now, there is an important caveat here, which is that I am working with a decomposition of the verb phrase and Champollion was assuming a single V head. For him the 'lexical' nature of event closure means that the event variable is bound before the argument DPs are introduced. In the Ramchandian approach to verbal lexical meaning, on the other hand, we have a rather decomposed event structure with arguments interleaved in specifier positions at various heights. Specifically, it means that introducing event closure at the edge of the first phase does not necessarily have the happy outcomes for quantification that are natural consequences of Champollion's original proposal. Nevertheless, I maintain that the idea here is essentially the same as Champollion's, but in a syntactically updated sense. To get the results concerning the interaction with quantificational arguments, we need to be more specific about how nominal arguments are merged during the course of the building up of the proposition. In brief, my assumption here is that nominal projections too are partitioned into a symbolic D_μ domain and a higher domain of instantiation, or reference. The part of the argument that is merged in the first phase is actually not a fully fledged referential projection or phase, but the lower portion of the nominal argument, that contributes its conceptual semantics to the build up of the D_μ domain before the deployment operator is merged. Only later on are these arguments given referential status, existentially bound, or quantified. Thus, all quantification lies systematically *outside* the level of event closure, as in Champollion's system. A detailed exposition of the interleaving of nominal and verbal functional sequences to build the proposition is however beyond the scope of this monograph, and will have to remain the major promissory note of this work.

So at the level of AspP we have a property of Relations that link the utterance context d with an existing event that is being demonstrated/described in d . That event has conceptual/perceptual properties as characterized by u . At this point temporal information can be added to the event description that was impossible before.

The quotational semantics outlined above will be used in what follows to implement the idea that there is a level of composition of concepts that is intuitively just about abstract properties, and that there is a level of composition where the information about the instantiation of these properties is expressed. Words themselves do not have truth conditions, they present ingredients to truth conditions. Truth conditions only arise when words are deployed in a context d , and with respect to which their particularity is established.

While the shift proposed here might seem drastic, I think it important to notice that the distinction between the domain of D_μ and the elements in it, and other primes of the syntactic computation which we will see when we reach the higher zones of the clause, closely parallels the distinction between lexical and functional items in classical generative grammatical analyses. The ‘lexical’ open class items are members of D_μ , and they consist of a triple which has some representation of conceptual content as its third member. I have written this conceptual content in standard lambda notation (although this might not turn out to be the best way to think about it in the end) in order to keep as much as possible to standard expectations of what the meaning contribution of a lexical item is, and to allow for integration into representations that feed truth conditions. In addition to the third member of the triple, which is something which I have elsewhere called ‘conceptual content’, I assume that there is also structural semantic content, which always exists whenever there is structure, but which will be the *only* thing that exists for functional elements. Functional elements will not be members of D_μ . In other words, they are not in the scope of the deployment operator. Speakers do not actively deploy functional elements in the same way that they active deploy lexical concepts to build content. Once we are in the higher reaches of the clause, we will not be building complex symbols, we will be deploying them in world, location and time, in an actual context and the denotations of the representations formed up by the computational system will look more like what we are used to. At that point, the formatives that are merged have a syntactic specification, a structural semantic denotation (and sometimes a phonology) but no conceptual content. So in this system, this is the way lexical vs. functional is captured, and it is also similar in spirit to the idea of asyntactic roots that is an important part of Distributed Morphology (DM). Unlike DM’s acategorial roots however, members of D_μ have syntactic decompositional information (in order to account

for selection and event structure generalizations), and more importantly, members of D_μ do not have to be atomic. Integrating the symbolic complex with contextual parameters to create an assertion bound in worlds and times is the job of the second phase of the clause, and it is with respect to this kind of instantiation-oriented syntax that members of D_μ are ‘innocent’.

The view of semantic layering I am building up here has much in common with the syntactic and semantic ideas being pursued by Wolfram Hinzen in recent work (Arsenijevic and Hinzen 2012, Hinzen and Sheehan 2011, 2015, Hinzen 2017). I share with him the belief that meaning is grammatically grounded, and that we need to build a new natural language ontology for semantics to match how language actually does the job (see also Moltmann 2017). But more specifically, I share with Hinzen the idea that the lexical symbols of language are primitive ‘essences’, and that the lower parts of the clause are then clothed with grammatical information to allow reference to specifics. The following quote comes from Sheehan and Hinzen (2011).

"In particular, while lexical items such as MAN or RUN reflect perceptually based conceptual classifications, and in this sense have a form of semantic reference, they are not used to refer to a particular man as opposed to another, or an event that happened yesterday over a certain period of time. ? Grammar based means of referring on the other hand, systematically establish relations of relative distance between the object of reference and the immediate features of the speech context. ..." Sheehan and Hinzen 2011, pg 2

When it comes to the notion of instantiation, Hinzen and Sheehan (2011) also argue that "Generally speaking, as we move from a given lexical root to the edge of the phase that it projects, reference becomes more specific ? Reference is in this sense an ‘edge phenomenon?’. " (pg 4) The role of deixis in establishing actual reference is also emphasized in that work.

The differences between my proposal and Hinzen’s framework lie in the fact that he divides the grammar of language into three separate ontological domains corresponding to objects, events and propositions, while I am assuming that events (situations in my terms) form an extended projection with the clause and ‘proposition?-denoting structures. Thus the analogy with the nominal domain works out a little differently in matters of detail than in his conception. In particular, the introduction of deictic information occurs quite low in my own structures, at the edge of vP (my EvtP). The Hinzen account also does not have an explicit domain of symbolic denotation, or a sharp ontological break at the little vP. Still, philosophically speaking, many of the substantive ideas behind the proposal I am arguing for are shared with the Hinzen approach.

1.3 The Grammar of Auxiliation

So far, I have laid out the intuitive ontological and formal background to the project stated in schematic terms. I have also outlined the general motivation from typology for making such a move. In the rest of the book, I will flesh out the details of the proposal within a particular empirical domain to give it substance and plausibility. The main testing ground for the proposal, and the data that will be used to ground many specific aspects of the theory, will be the properties of the English auxiliary system. I will attempt to account for the core semantic and ordering properties in this domain.

1.3.1 Ordering

As is well known, the ordering of the English auxiliaries is rigid (cf. Chomsky 1957), as illustrated in (14).

(14) (a) {T, Mod} < Perf < Prog < Pass < V
(b) He could have been being interviewed.
(c) *John is having returned.
(d) *John is being hunting.
(e) *John seems to have had already eaten.

Most modern syntactic representations of the phrase structure of the English verbal extended projection simply assume a templatic ordering of Perf over Prog over Pass (Bjorkmann 2011, Sailor 2012, Aelbrecht and Harwood 2012, Bošković 2013), when these elements need to be explicitly represented. Linguists differ with respect to whether they simply represent Perf, Prog and Voice as functional heads (Bjorkmann 2011 and Sailor 2012) and handle the inflectional facts via ‘affix lowering’ or AGREE, or whether they in addition assume separate functional heads hosting *-en* and *-ing* (Bošković 2013 and Harwood 2013).

Within Minimalism, the assumption seems to be that some kind of selection is at work, and does not represent a universal functional sequence, and these projections are left out even for English when the literal perfect or progressive forms are not expressed in the sentence. But the account as it stands barely rises above the level of description, since the labels for the functional projections Prog and Perf are tailor-made for just progressive and perfect respectively, and no attempt is made at a higher level analysis or generality for their function. Thus, the deep questions about what is responsible for this rigid ordering, are never even asked in a meaningful way; they are essentially sidestepped by the stipulation of a deliberately locally descriptive template.

In fact, I think that the phenomenon of auxiliation is an interesting one in its own right. While many languages express tense, modal and aspectual notions via inflectional morphology, the auxiliating languages take a more isolating strategy. If were to tackle the typological generalizations about order of composition from a morphological perspective, we would give ourselves the additional theoretical/architectural question of the precise relationship between syntactic hierarchy and morphological order. While much is known in this area empirically (see specifically the Mirror Principle of Baker 1985 and subsequent work), it introduces an extra layer of theorizing to any discussion of the compositional semantic problem. With respect to isolating (and in particular auxiliating languages), the semantic and hierarchical issues remain the same, but some of the worries about the internal structure of words and their relationship to the syntax can be sidestepped. It is for this reason that the English case is an interesting one to solve. It cannot be relegated to the padded cell of morphology (for those who think morphology *is* that kind of encapsulated world). It *must* be dealt with in the phrase structure and in the interpretation of phrase structures.

In Ramchand and Svenonius (2014), we argue that the cartographic orderings observed deserve an explanation in terms of the semantics that underwrites the building up of clausal semantics. There have been attempts to explain these orderings in semantic terms in the previous literature (e.g Schachter 1983), but not in a convincingly general way. One of the main goals of this short monograph is to take the relatively concrete problem of auxiliation in English and show how a different kind of account can be crafted when brought together with an explicit set of proposals about the semantics of verbal meaning. The idea is to make good on the promissory explorations of Ramchand and Svenonius (2014) and provide a detailed and explicit exposition of the auxiliary system of English in terms of structured situations that begins to make predictions for other languages. In other words, I will attempt to provide a compositional semantic theory that exploits the notion of distinct domains of composition from a semantic point of view.

1.3.2 Lexical Specification and Polysemy

I stated earlier that the methodology I am going to use will be somewhat inverted from the standard procedure. One reason for the difference is that a guiding motivation for the system will be for individual lexical items to be as unified as possible. In other words, I am going to try to assume polysemy by underspecification to as great a degree as possible.

To give an example, suppose we have two meanings corresponding to English /bæjk/—‘financial institution’ and ‘ground sloping up from a river’. Ev-

erybody is happy in this case to assume that we have *bank*₁ and *bank*₂ in our lexicon, which happen to accidentally share a pronunciation. This is the standard case of ambiguity, or homonymy. Equally standardly, we are less inclined to see the uses of *game* in the (a) and (b) sentences below as two separate lexical items.

(15) (a) John is playing a game of solitaire.
 (b) I don't think Mary is serious—she's just playing games with me.

These are cases of lexical vagueness, or polysemy.

When it comes to verbal meaning, things become a little bit more tricky. Are the two uses of *break* in (16) the same lexical entry, or two different but 'related' ones?

(16) (a) John broke the stick.
 (b) The stick broke.

While the problems have not been solved in all cases, they have been noticed and discussed extensively in the literature on lexical meaning.

Lacking so far however, is a serious discussion of the polysemy of certain functional items and functional morphemes like participial endings and auxiliaries. There are good reasons for this. In the theories of morphology that actively interact with and engage with syntax, functional items are often assumed to be essentially devoid of interesting conceptual content, outside of the syntactically active features they possess.

The functional polysemies that I will be concerned about concern the interpretation of the *-ing* participle in English, the *-ed* participle, and the auxiliaries *have* and *be*. I will insist on meanings (however abstract) for all of these formatives, and will consider it a virtue if any analysis can run based on fairly unified denotations. I list the well known English polysemies in this domain below for convenience.

(17) **The *-ing* Participle**

| | |
|-------------------------------------|--|
| (a) John is running. | <i>Progressive -ing: activities</i> |
| (b) John is drawing a circle. | <i>Progressive -ing: accomplishments</i> |
| (c) The dancing children are happy. | <i>Attributive -ing participle</i> |
| (d) Dancing is fun. | <i>Gerundive</i> ⁸ |

⁸ The gerundive here actually encompasses a host of subtypes, as is well known from the literature. In fact, since this book deals with verbal extended projections, it will not deal with these in any detail, although it will try to provide a motivation for the existence of this family of more nominal like expressions built around *-ing* forms.

(18) **The *-ed* Participle**

- (a) I have rejected that idea. *Perfect participle*
- (b) The offer was rejected. *Passive participle*
- (c) The rejected offer. *Attributive -ed participle*

(19) **Auxiliary and Main Verb *Have***

- (a) John *had* a heart attack. *Light verb have*
- (b) I *have* a brother. *Possessive have*
- (c) I *have* seen that movie. *Perfect have: experiential*
- (d) I *have* broken my arm. *Perfect have: resultative*
- (e) I *have* lived in Paris a long time. *Perfect have: universal*

(20) **Auxiliary and Main Verb *Be***

- (a) John *was* in the garden. *PP-predication be*
- (b) The computer *was* broken. *AP predication be*
- (c) The metal *was* hammered flat. *Passive be*
- (d) The thief *was* running. *Progressive be*

Since I will be concentrating on auxiliary constructions, I will not be attempting to unify the *-ing* participle with its gerundive, more nominal like distributions, although I will pave the way for a potential unification. In the case of the auxiliary *have*, I will not attempt to unify the lexical domain *have* with the functional domain *have* (see Myler (2017) for an attempt in this direction).

Many researchers think of *be* as an aspectual auxiliary and/or the mere spell out of tense and so do not regard it as a desideratum to give a unified entry for such ‘functional’ items. However, even functional items in this sense are elements of the lexical inventory. To the extent that all the verbs pronounced in the same way and sharing a morphological paradigm have the ‘same’ entry, I consider that an advantage. Recent work by Myler attempts to give a unified account of *have* across its main verb and causative uses. This is done by completely voiding *have* of any semantics whatsoever, while listing a set of insertion contexts. However, unless the insertion contexts themselves ‘have something in common’ as opposed to just being a disjunctive list, this is also not a real unification. Throughout this monograph, the focus will be on the reusable linguistic ingredients of the system and how they are efficiently deployed in the recursive combinatorics in building propositions. Reusability (and by extension polysemy) will be seen as a design feature of the system rather than a bug. As a methodological principle then, I will be guided by the fact of polysemy and seek analytic unities to underlie them.

In addition to the aspectual auxiliaries, the modals in English will also be given denotations, and here we are on more familiar ground when it comes to underspecification and polysemy. The early Kratzer position (Kratzer 1977) is

to give a simple underspecified semantics for each modal and allow the richness of the meaning to emerge from the interaction of that meaning with contextual and pragmatic information. Hacquard (2006) goes one step further and allows the meaning of the modal to be affected by the syntactic height at which it is merged. My own account will continue the Hacquardian line of thought and attempt to fill in an underspecified modal semantics compositionally via the nature of the prejacent that it combines with.

(21) **Modal Polysemies (eg. Circumstantial vs. Epistemic)**

| | |
|--|-----------------------|
| (a) John <i>must</i> be in his office now. | <i>Epistemic must</i> |
| (b) Mary <i>must</i> pass that exam. | <i>Deontic must</i> |

Across the board, I will be seeking a unified an underspecified semantics for the ingredients of auxiliation. A unified underspecified semantics has the virtue of being simple, easy to acquire and part of a modular system of interaction with other elements. Underspecified meanings will interact both with structure, and with contextual information in order to get the correct effects on the ultimate truth conditions of those forms.

Polysemy is a pervasive feature of human language and affects lexical verbs as well as functional items. This is in fact the standard case. Moreover, it is clear from the processing and psycholinguistic literature that greatly polysemous items are actually quicker to access and process than highly particular items (cf. what Harald Baayen calls ‘contextual diversity’ Adelman et al. 2006). While homonyms have a inhibitory effect on processing polysemes seem to facilitate each other. Any verbal item can be deployed by a speaker to describe a wide range of actual event particulars in practice because each property encoded by the verbal element is *partial* and can be deployed by the speaker in potentially many ways.

Consider again, the core idea behind the quotational semantic approach, repeated here in (22).

(22) **EVT** : $\lambda u \lambda d \lambda e [\text{UTTERANCE}(d) \wedge \text{TH}_\delta(d) = u \wedge \text{CONVEY}(d, e)]$

The deployment relation at the heart of the quotational semantics I will be proposing could correspond in context to anything from full to partial imitation of certain properties of the event, invocation of a *particular subset* of the conceptual properties of the linguistic item, or even invocation of event properties by deploying a saliently non-descriptive or oppositely descriptive linguistic item. All we need minimally is the quite loose CONVEY relation which in a successful communicative context could even have the following content as a special case: ‘e makes me think of this property even though it

does not literally have it'. I suggest that the contextual relation at the heart of the deployment of the lexical item u in the utterance situation d to describe e is what makes sense of ideophones, metaphorical uses, sarcasm and makes itself felt in the general pervasiveness of polysemy for lexical items.

As we have seen however, in the standard literal case, CONVEYing and deployment ($\text{Th}_\delta(d)$) reduce to their most standard incarnation, which is that the lexical item denotes a property of events that the described event has, namely (23).

$$(23) \quad \text{EvtP} : \lambda d \lambda e [\text{UTTERANCE}(d) \wedge \sqsubset u \sqsubset (e) \wedge \text{DESCRIBE}(d, e)]$$

This is what I will use in the representations in future, putting aside a detailed exploration of metaphor and sarcasm and other metalinguistic like uses of those lexical items for future work.

1.4 Morphology and Spanning

On the morphosyntactic side, this book will adopt a somewhat non-standard, but I hope fairly transparent view of the syntax morphology interface and its relation to lexicalization.

$$(24) \quad \text{Span:}$$

A span is a contiguous sequence of heads in a complementation relation.

As noted already in the discussion of the phenomenon of auxiliation itself, English is an example of a language where there is very little agglutinative morphology, and is primarily analytic. This means that we will not have to take a strong stand on the relationship between syntactic dominance and morpheme order and how it is achieved. English is also a language which routinely orders specifiers before head and heads before complements, transparently reflecting the mapping to linear order legislated by the LCA (Kayne 1994). This means that with respect to linear order also, I will simply assume that precedence directly reflects syntactic height, and will not have to engage with any word order movements or fancy linearization algorithms. Luckily, therefore, I will not have to take a strong stand on many of the syntactic issues that are at the forefront of current debate.

I will primarily be concerned here with cartography in the extended sense, i.e. with the syntactic features that correspond to the functional sequence of the clause (regardless of whether this turns out to be spare and minimal, or somewhat richer than the labels used in the classical GB period. These are all 'category' features, and they form complementation relationships. As dis-

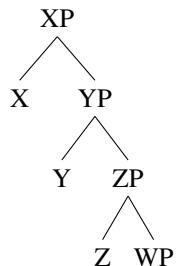
cussed extensively in the literature, there are well known strong typologically supported effects of syntactic structure on morphology (Baker 1985, Julien 2000, Cinque 1999). The standard approach to such ‘mirror’ effects is head movement, a device currently deprecated in minimalist syntactic theorizing. Brody (2000) offers a different view based on direct linearization, whereby the heads in a complementation structure are linearized together as a word in reverse order (i.e. dominance in the complementation line corresponds to subsequence in terms of the linear order of the morphemes within the word). I will be adopting a ‘direct linearization’ view of the effects of head movement in this monograph.

The importance of the functional sequence and the somewhat more fine grained view of the phrase structure that I will assume means that there are a couple of core issues where my implementation will give rise to theoretical choices. English is not well endowed with morphology in the first place, as we have said, but there are going to be clear cases where the syntactic properties of a lexical item in my analysis are related to the syntactic features on adjacent syntactic heads. These are the situations that would classically either require head movement, or morphological ‘fusion’ before vocabulary insertion, if I were to adopt more standard models. For example, Distributed Morphology (henceforth DM) requires lexical items to be inserted under a single terminal node (for an explication of the core properties of DM see Halle and Marantz 1993, Marantz 1997, Embick and Noyer 2001, Harley and Noyer 1999). In cases, where this is not sufficient, various devices such as ‘fusion’ or ‘fission’ are required to feed vocabulary insertion. Instead of either head movement or fusion, I will be adopting the implementation that is known as ‘spanning’ (after Williams 2003). In the spanning view of things, a morpheme may spell out any number of heads in a complement sequence (see Ramchand 2008, Adger et al. 2009, Caha 2009). Spanning itself is similar in some respects to the system in Brody (2000) described above, in sharing a direct linearization solution to the problem of ‘head movement’, but different in the sense that it is not confined to the spell out of individual morphemes related to single heads—it operates on morphologically holistic forms which are specified for more than one category feature (see Adger 2010, Svenonius 2012 and Bye and Svenonius 2012 for a discussion of the system and how it differs from the original Brodian conception.)

Although I will differ from DM in recognizing SPANS as an input to spell-out, like them, I will assume a model of late insertion, whereby the lexical exponents are selected after the syntax in the relevant domain is complete.

Thus, in an abstract tree structure such as the one shown in (25), vocabulary items can be specified via their category features, to spell out any contiguous span of heads in the complement sequence.

(25)



In this system, lexical items such as $LI < X >$, $LI < Y >$, and $LI < Z >$ are possible in addition to the ‘spanning’ lexical items $LI < X,Y >$, $LI < Y,Z >$, and $LI < X,Y,Z >$. Crucially, the lexical item $LI < X, Z >$ will be uninsertable since it violates a requirement on the *contiguity* of lexical items. In addition, I will assume the ‘spanning’ version of the elsewhere condition here which says that while an LI can be inserted into a structure that contains its category features as a subset, it cannot be inserted in a tree where the LI does not possess the feature.⁹ If we tried to insert $LI < X,Y >$ to span category heads X, Y and Z, then *specificity* will be violated. These two principles are listed below in (26).

(26) (a) *Contiguity*: An LI can only be inserted in a phrase structure tree as the exponent of a SPAN if it spans a contiguous sequence of heads in the structure.
 (b) *Specificity*: An LI can only be inserted in a phrase structure tree as the exponent of a SPAN if its lexical entry contains all the features in the span.

In the current implementation, we move away from the late insertion model of expressing the separation between conceptual content and syntactic structure. Instead, the ‘separation’ is enforced via the fact of separate components within the lexical item itself. The lexical items are then combined directly in the first phase via Merge, contributing in some cases larger syntactic spines

⁹ See Caha (2009) for an explication of the equivalence between DM’s version of Elsewhere in terms of underspecification (‘The Subset Principle’) and the version required by the spanning approach (‘The Superset Principle’).

to the whole, as in the spanning intuition. Generalizations over phrase structure hierarchy must now be stated as local generalizations about heads within lexical items, or as macro generalizations derived from the universal compositional principles embodied in the QQS. This immediately deconstructs the fseq in its templatic incarnation and gives us the challenge of recouping its effects from other sources, continuing on from the agenda in Ramchand and Svenonius (2014). In the direct merge view, we no longer have competition for insertion, or an elsewhere condition. Elements of μ will contribute the syntactic features they have and no more. Unprojected features however, will be possible under limited conditions.

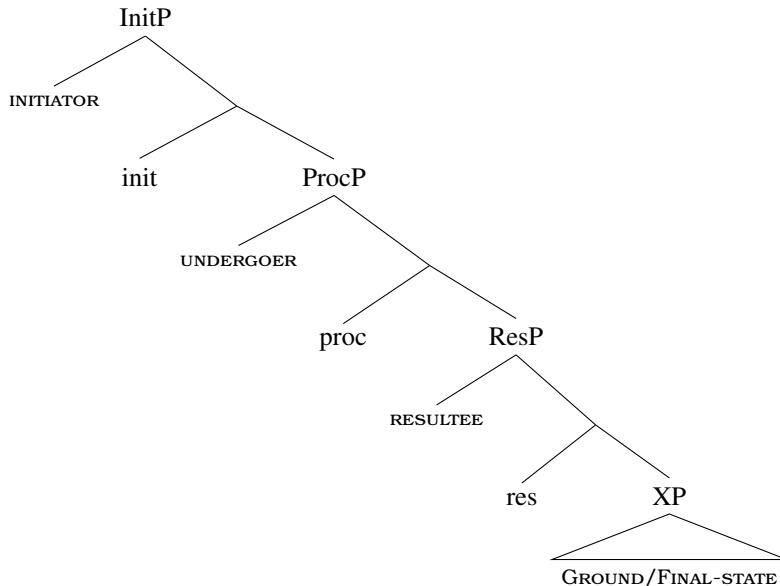
Note that this system differs in many ways from standard DM. Apart from the abandonment of late insertion to express ‘separation’, LIs as we have seen, are explicitly specified for category features, including LIs that DM would consider to be acategorial roots as well as more ‘functional’ items. This goes hand in hand with the abandonment of the late insertion model, which is no longer strictly necessary to the separation intuition. However, the system still has much in common with general constructional approaches to argument structure in that it does not contain a generative Lexicon in any way, and does not overtly express argument relationships in the Lexicon. One single generative component located within the narrow language faculty is what is responsible for the productive generation of syntactic and semantic structures. All possible argument relationships are indirectly specified via the category features which built events of a particular type, but are not directly represented as argument structure frames (see Ramchand 2008 for discussion). In many ways, given the introduction of the D_μ domain for the first phase, there is a sense in which the first phase really is a kind of Lexical Syntax, to use the term originally favoured by Hale and Keyser in their early ground breaking work in this area (Hale and Keyser 1993b, Hale and Keyser 2002).

For concreteness, I lay out the framework of Ramchand (2008) for expressing verbal meanings to illustrate the decomposition of event structure I will be assuming in this work.

I have argued in that work, and also in Ramchand (2016) that there are clear generalizations across languages with respect to *aktskript* and argument expression which converge on the following facts: verbal and argumental items that describe or undergo the result of an event are hierarchically embedded under verbal and argumental items that describe or undergo change simpliciter; these two in turn are hierarchically embedded under any verbal and argumental items that express static or dynamic causes of those changes. One could represent these relationships in a template as we do for in the rest of the func-

tional sequence, as follows.¹⁰ As in Hale and Keyser, the causing or ‘leads-to’ relation between subevents corresponds to hierarchical embedding.

(27)



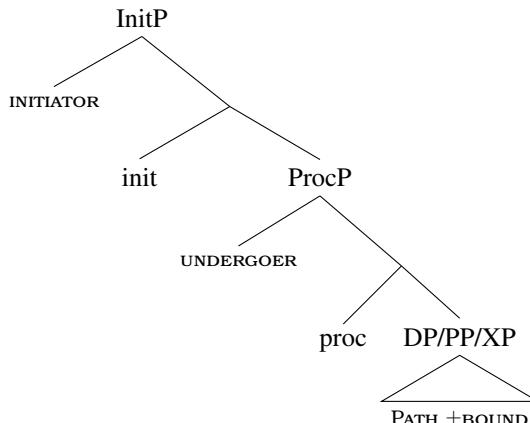
The above tree expresses the maximally expanded subevental structure for caused changes leading to a result, with a stative predication embedding a dynamic one, and the dynamic one in turn embedding a stative one. Thematic roles do not need to be listed separately, nor do their properties need to be memorized or known in advance. Interpreting phrasal embedding as causation will ensure the relative prominence of the different argument positions, and the minimal relationships of property-holding (both static and dynamic) will derive specific entailments for the different positions.

¹⁰ In fact, in previous work (Ramchand 2012; Ramchand 2016) I have argued that there are aspects of this template that should need to be stipulated, but which follow from the recursive deployment of a number of primitive relations that operate over event properties as a matter of human cognition. Subevental embedding corresponds as a matter of general principle to the cause/leads-to relation. I propose to limit recursion to structures with a maximum of one dynamic predication per event phase. This is a constraint that comes from our general cognitive relationship to event perception— independently perceived dynamic change corresponds in interpretation to a *distinct* event. To be parts of the ‘same’ event, there can only be one independently represented dynamic core. Finally, the thematic roles are restricted to the holding of either a static property, or of a changing property. They do not need to be ‘selected’ but are introduced via predication at each level of subevental description. These are abstract properties of the cognitive glue that puts events and their participants together, and they are patterns in the abstract system of grid lines that serves to underwrite and organize the verbal lexical labels we then learn.

The highest specifier is the ‘holder’ of a property which ‘leads-to’ the change occurring. This is just a fancy way of saying INITIATOR. The middle specifier is the ‘holder of a changing property.’ This is just an UNDERGOER. The lower predication expresses a property that comes into being/ is caused or led to by the central dynamic event. It is thus a ‘result’ and the ‘holder’ of that result property is the ‘holder of result’, or RESULTEE. The labels on the tree should therefore be seen not as labels in a template, they are there for ease of readability. The functional sequence here is actually quite spare, once the effects of hierarchy and predication are factored out.

In addition to the maximal subevental expression above, activities and accomplishments can be built from structures that lack the lowest result projection. Bounded paths give rise to verb phrases that are classified as accomplishments in the literature, while unbounded paths give rise to activities.

(28) **Activities (Path –bound) and Accomplishments (Path +bound)**

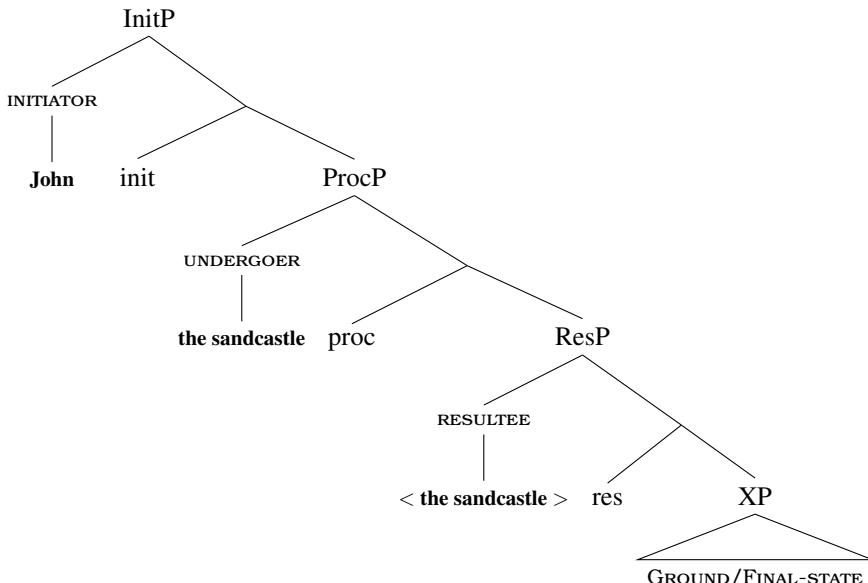


In this system, the event structure hierarchies and participant relation hierarchies track each other quite directly, and follow from a single decompositional structure.

But how does this structure relate to actual lexical items? The answer depends on what one’s assumptions about lexicalization are, and there are many possibilities here. The traditional view, continued in DM, is that lexical items attach by insertion under *terminal nodes* of the syntactic representation. However, this is in tension with the increased elaboration of phrase structure (necessary to capture generalizations about hierarchical structuring), where devices such as head movement, morphological fusion and allomorphic selection have been employed to capture the fact that a single ‘word’ seems to express a composite of syntactic information that is ranged hierarchically.

Thus, in discussing the lexicalization of structure, I will assume a model which lexical verbs come listed with category features that express their insertion possibilities, in terms of spans. The English verb *destroy*, having all three features *init*, *proc* and *res* (or possibly just $V1_{state}$, $V2_{dyn}$ and $V3_{state}$) identifies the full structure ‘synthetically’.

(29) John destroyed the sandcastle



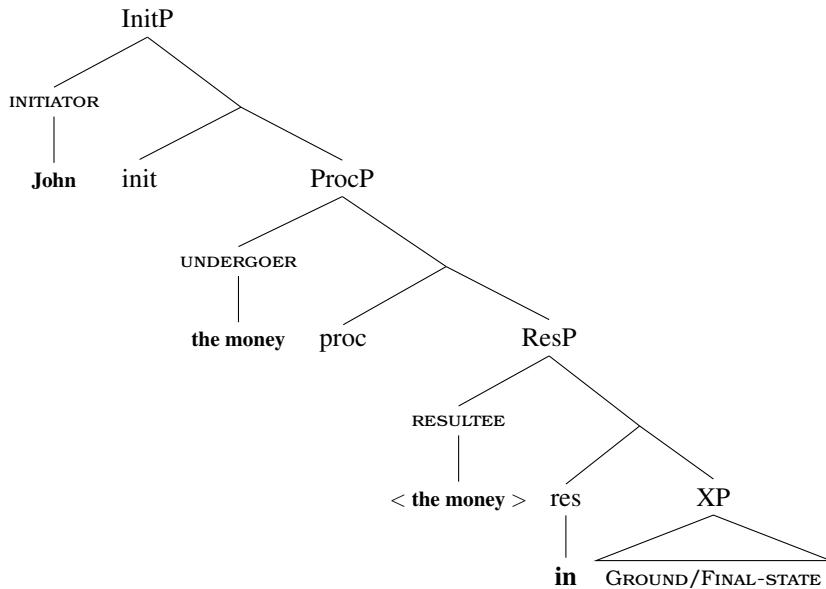
In the present implementation, we deal directly with elements of the D_μ domain. Syntactic structure is thus introduced by the merge of the elements of D_μ , where the syntactic component consists of a single contiguous spine of categories (to account for the spanning intuition). The lexical item also contributes its conceptual content which is unified with the structural semantic contribution of the node(s) in question.¹¹

Allowing for ‘spans’ (cf. also Williams 2003) in this way, we are in a position to see that languages lexicalize these structures in a variety of different ways, depending on the inventory of the building blocks at their disposal. In English, we also find a more analytic version of this construction, where a

¹¹ The Nanosyntactic approach of Caha (2009) recognises non-terminal spell out, and also recognises ‘treelets’ as the syntactic part of the lexical item’s specification. However, in that system, trees are not built by the direct merge of lexical items themselves, but by a rewriting process based on ‘match’. The nanosyntactic approach further differs from the present one in that syntactic ‘movements’ can be triggered in order to feed lexicalization under ‘match’.

particle explicitly identifies the result and combines with a verb that does not usually license a direct object to create a derived accomplishment structure with an ‘unselected object’ (Simpson 1983, Carrier and Randall 1992).

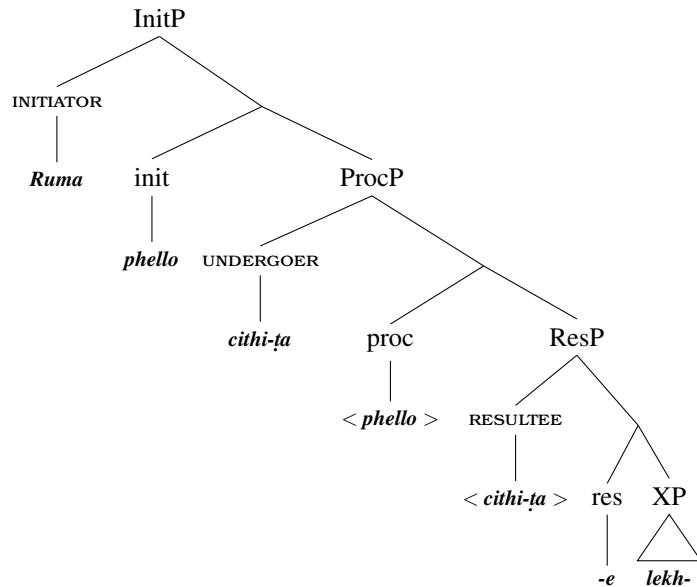
(30) John handed in the money.



We can compare this with Bengali, which has an analytic construction: the perfective participle *lekh-e*-‘written’ identifies the *res* head, while the ‘light’ verb *phæla*-‘drop/throw’ lexicalizes *init* and *proc*.

(31) Ruma cithi-**ta** **lekh-e** **phello**

Ruma letter-DEF write-PERFPART drop/throw-3RDPAST
 ‘Ruma wrote the letter completely.’



Bengali is of course a head final language. Quite systematically, aspect appears outside of the main verb stem, and tense in turn appears outside of that. They then line up sentence finally as V-Asp-T. We can remain agnostic here about how that word order is derived, but note crucially that the ‘higher’ functions of process and initiation in the verbal decomposition appear to the right of the ‘lower’ description of the result state (the participle). This is exactly the order you would expect from a head final language with this proposed hierarchical structure.

In previous treatments, complex predicates such as this, and even the English particle verb construction have posed paradoxes for lexicalist theories of argument structure. On the one hand, they are clearly morphemically compositional, and it can be shown that the component parts are even independent syntactic units. On the other hand, the combination of lexemes changes the argument structure properties (something that lexicalists assume to be in the domain of the lexical module) and the constructions are monoclausal by all diagnostics. The constructivist view proposed here accounts for the predicational unity of the complex predicates as well as their resultative semantics. The complex predicate construction of the resultative type, the verb-particle constructions and the synthetic English verb ‘destroy’ have essentially the same hierarchically organized components, but are just lexicalized/linearized differently.

In all of the above examples, it is still possible to conceive of lexical insertion in a more traditional manner under terminal nodes, with head-to-head

movement in the syntax, or in the morphology as the need arises. I present the multiassociational/spanning view here because I believe it requires fewer ancillary ‘modules’ (such as ‘Fusion’ in the morphology), and because it highlights the sharp difference between conceptual content and structural semantics.

What does it mean for a lexical item to have a particular syntactic feature label in this event decomposition? It means simply that it is an element of μ that contributes a sensory/perceptual or cognitive property of that subevent. Thus, a composite verbal item like *destroy* potentially contributes conceptual content to all three subevents in its syntactic specification, although there may be variations in how rich and specific those properties are from verb to verb. In the Bengali construction, the conceptual content associated with the initiating event is extremely abstract but the specification of the result state is rather rich. The opposite is true in the particle construction in Germanic, which the rich lexical content resides in the initiation and process while the result is rather abstractly encoded. In general, I assume that all of these conceptual contributions to the specification of the event property are *unified*, added together to form the final composite description.

I will follow my own previous work in assuming that it is in principle possible for syntactic information to remain unprojected. However the possibility is highly constrained in the sense that any non-projected category feature in an element of μ must in the normal case be licensed in an agreement relation with another μ that bears that feature. Any two elements of μ whose category features are in an Agree relation must moreover unify their semantic content. The principle is given below in (32).

(32) *Constraints on Projection:*

- (a) Non-projection of category features of an ordinary member of μ is in principle possible, constrained by Agree.
- (b) Agreeing categorial features must unify their conceptual content.

Finally, and crucially for what will follow, I will assume one further principle on the regulation of lexical insertion that is familiar from the literature, but with a twist that relates to the other assumptions followed here. I will assume a pervasive system of Blocking that adjudicates different possible lexicalizations of structure.

(33) *Poser Blocking*¹²

The lexicalization of a SPAN by a single element of μ is always chosen over lexicalization of an identical SPAN via separate elements of μ .

The reason we have had to be so clear about the principles of lexicalization is that the phenomenon of auxiliation raises some of the very same issues that are raised by complex predicates. I will be thinking of auxiliation as the process by which structures that *could* in principle be lexicalized by single inflected verb forms can also be lexicalized piecemeal by individual forms. In doing so, only one of the forms will be finite, while only one will contribute the main event properties to the situational description. Auxiliary structures are a species of analytic spell out for the same abstract hierarchical structures that could in principle be spelled out synthetically. Auxiliary structures are therefore a crucial part of the puzzle concerning what these pieces/ingredients of verbal structure are.

Phrasal blocking is going to play a starring role in much of the discussion in the analysis of *en/ed* and *-ing* in English and their relationship to the corresponding simple verb expressions. It will turn out to be pervasive, and it will allow us a simpler statement of the facts with fewer item specific stipulations concerning selection. Poser blocking and its role in participle constructions is also anticipated strongly in the work of Lundquist (2008), to whom the ideas behind the treatment in chapter 3 are indebted.

1.5 Roadmap

The core idea is that the quotational semantics approach makes it possible to distinguish between a concept building layer of the clause before the event variable is existentially bound, and the properties that arise in the context of instantiation in world and time. At this level, the proposal is that elements of type μ are restricted to nontemporal, nonmodal properties of events, learned by implicit generalization over experience. Following Champollion, existential closure happens at the edge of the first phase. Thereafter, event-related aspectual worldly and temporal information can be added as properties of the

¹² Poser blocking is so-called after the first close examination of cases where the expression in terms of a single form blocks the expression in terms of multiple words. This kind of blocking is effectively disallowed in a system like DM where the competition for lexical insertion plays out only in the domain of single terminals. The operations of fusion and selection which implicate adjacent heads can allow the result of competition to include information from directly adjacent heads in the phrase structure, but no further. See Embick and Marantz (2008) for a recent discussion. The grounding assumptions of DM are so different from my own in this monograph that it is not immediately obvious to me whether the generalizations that underlie the use of the principle in chapters 2 and 3 can or cannot be captured in a DM framework using different principles.

relation between the utterance eventuality and the eventuality being described. These assumptions work together to drive ordering effects in the verbal domain.

In many ways, I see this monograph as an attempt to reconcile the important and robust results that have emerged from the syntactic tradition (particularly cartography) with the vocabulary of semanticists. More and more it seems to me that semanticists and syntacticians are not part of the same conversation. More and more the results of one subfield seem irrelevant to the results of the other. The ontologies are not commensurate, and the questions addressed are often not the same. Furthermore it is even technically difficult to reconcile terminological assumptions and assumptions of what the ‘primes’ of the representations are in order to get a conversation off the ground. In turn, we in formal linguistics are not in a position to present a useable set of hypotheses and predictions to psycho- and neuro-linguists because of a lack of precision about the division of labour between conceptual content triggered by lexical open class items and linguistic structural semantic content. Even if the implementations offered in this monograph turn out to have more elegant alternatives. I hope at least that they are concrete enough to be stimulating, and that they will generate different and productive ways for semanticists and syntacticians to look at the central issues treated here.

In chapter two I look carefully at the English progressive, arguing for its position within the first syntactic phase of the clause, and within the domain of event concept building. One of the key claims of this chapter is that an explicitly intensional or ‘modal’ account of the progressive in terms of possible worlds is both undesirable and inadequate empirically. Instead, I will argue for a more primitive relationship between events and event parts as part of the denotation of the *-ing* morpheme.

In chapters three and four, I look at the interpretation and distribution of the *-ed* participle in English which appears in both the passive and perfect constructions in English. I will argue that while the participle is built within the first phase, the *have* auxiliary of the perfect is in the domain that is entered *after* the existential closure of the event variable and where temporal properties become possible to state. This will account for the ordering with the progressive. The denotation proposed for *have* will be used in analysing the difference between the perfect and passive constructions in that that latter will be built using the auxiliary *be* which will be argued to be generated low, as with the progressive.

Chapters five and six deal with modals and modal ambiguities. Modals will be situational modifiers operating at different heights. This investigation will

also take us into the higher reaches of the clause where situations will be anchored to the contextual utterance information.

Chapter seven is the epilogue in which we return finally to the problem that started us off on this journey in the first place—the rigidity of auxiliary ordering in English. In this chapter I take stock of the analyses proposed in the monograph and discuss how the system proposed actually forces the order we find, and disallows the others. I assess the principles and assumptions that were used to achieve these effects. I also briefly discuss the implications of the proposed system (which is in fact quite general) for other languages and constructions, offering speculation and suggestions for further research.

Ideally, such a big change in proposed semantic primitives should be justified on the basis of many different languages, but in the context of a single volume I will have to content myself with a spread of data within just one language. Even this is a mammoth task. Each of the constructions of English I will examine has had many books written about it alone. The advantage of this situation for the present work, is that quite a lot is now known about the syntax and semantics of these constructions, and much can be built on with confidence. The groundwork has been done in many cases in this well researched area, and I am in the enviable position of being able to build on it in spelling out the shift in perspective. Research such as this within a cumulative endeavour has thus two distinct aspects to it: one is to account for the insights and successful generalizations that are currently captured in the literature, and the second is to attempt resolutions of some of the problems and paradoxes that still remain.

2 The Progressive in English

I start my analysis of the English auxiliary system with the progressive. There is a good syntactic reason to treat the progressive before the perfect or the modals, since it occurs lower than them in the structure. There is also a semantic reason. Building on the arguments of the introductory chapter, the progressive is going to provide the clearest linguistic evidence for manipulations in D_μ zone of the clause.

There has been a great deal of semantic work in the literature treating the semantics of the progressive, specifically on the difficulty of stating truth conditions for it when applied to accomplishment predicates. This problem has gone under the label of ‘the imperfective paradox’ in the literature (Dowty 1979), and is illustrated briefly below.

- (1) (a) Mary was crossing the street.
(b) Mary crossed the street.

The problem essentially is that if you try to give truth conditions for the progressive in terms of truth conditions for the bare untensed version of the event description ‘cross the street’, then any extensional version of event semantics is going to end up committing you to the street getting crossed at some time in the real future. However, it is quite obvious to speakers of English that an utterance of *Mary is crossing the street* does not in fact commit them to the assertion that she will ever make it to the other side. The problem is a fairly deep one for the way the semantics is set up: in order to precisely express the truth conditions of ‘cross the street’ we are forced to be explicit about the telos; in order to precisely express the truth conditions of ‘is crossing the street’ we need to get rid of the telos *but still describe an event that would have that particular telos if it did culminate*.

It seems like a paradox, we do not want *is crossing the street* to be true of ‘dancings’ or ‘walkings on the roadway’. We need to build in the semantics of street-crossing specifically, and that seems to require a description in terms of

the telos being attained. The by now standard solution to the problem has been to describe the attainment of the telos in terms of non-actual worlds (inertial worlds according to Dowty 1979, or continuation branches according to Landman 1992) in the semantics, while the assertion only applies to the world or situation that is actual. The accepted analysis of the progressive is essentially then, a modalized semantics, requiring some kind of notion of possible worlds in its definition.

If this is true, it has immediate consequences for the compositional semantic system that I am proposing in this book. This is because it essentially forces the progressive to be expressed at the level of situational particulars, where worlds can be quantified over. However, as I will show in the next sections, there is good evidence from the syntax of the progressive that it lies quite deep in the clausal structure, *within* the very first phase of ‘event building’. If we take the converging evidence from syntax seriously, then we would be forced to say that quantification over possible worlds/situations is possible in the very lowest parts of the clause. But if this is true, then why does something like the progressive occur in the zone below tense anchoring and modal modification by auxiliaries? As we have seen, the zone closest to the root appears to deal with event properties that are abstractions over world and temporal instantiations as a robust crosslinguistic fact, and the English progressive falls squarely within this zone. In the second section of this chapter, I will review the reasons for the modal analysis of the progressive in detail and conclude that it is neither adequate nor necessary for capturing the meaning of the progressive. I will instead propose a version which is most similar in spirit to Parsons (1990), but which avoids the notorious ‘part-of’ objections to that view. The account I will propose will be one in which the progressive creates a derived event property within the D_μ denoting zone of the clause.

2.1 The Syntax of the Progressive: The Progressive is in the First Phase

I wish to start however with the syntax of the actual progressive construction, and present some evidence for where the different components of it line up with respect to the more established elements of phrase structure such as V and v.

In what follows, I will show there is an important syntactic and semantic joint between progressive and perfect in English that should be represented explicitly by an abstract cut-off point in the phrase structure. Specifically, with respect to a number of different linguistic tests, the progressive, unlike the perfect, appears to pattern with the main verb and its arguments. The following three syntactic arguments were also presented in Ramchand and Svenonius

(2014) and Ramchand (2017), and the exposition is very similar to that found in the latter version.

2.1.1 Expletive Associates

In this subsection, I show data from Harwood (2011), Harwood (2014) who has argued that the progressive must be inside what is usually thought of as the vP phase. Harwood's evidence includes an extended argument based on classical VP ellipsis, and the idea that ellipsis is always targets a phasal spell-out domain. Although I am not equally convinced that traditional VP ellipsis is *directly* sensitive to zones the way Harwood suggests, his data from expletive associate placement is in some respects even clearer (and doesn't require the notion of 'flexible' phase that the full spread of ellipsis data requires). Part of the discrepancy between our accounts is that I take the semantic characterization of the lowest zone as lexical 'event description' as primary and axiomatic.

The data that follows however is relevant to both accounts. Harwood (2011) notes that the thematic subject of a verb in the expletive *there*-construction in English remains low in the clause and is confined to positions left-adjacent either to the main verb, or to the passive or progressive participles. It can never surface to the left of the perfect participle.

The examples in (2) with the full complement of possible auxiliaries, show that there is only one position in the sequence for an expletive associate, between Perf *-en* and Prog *-ing* (cf. Harwood 2011).

(2) (a) *There could have been being a truck loaded.
(b) There could have been a truck being loaded.
(c) *There could have a truck been being loaded.
(d) *There could a truck have been being loaded.
(e) *There a truck could have been being loaded.
(f) A truck could have been being loaded.

Even when the progressive itself is not present, we see that the position to the left of the perfect participle is still unavailable, while the position to the left of the main verb and passive participle is fine, as we see in (3).

(3) (a) There could have been a truck loaded.
(b) *There could have a truck been loaded.
(c) *There could a truck have been loaded.
(d) A truck could have been loaded.

Similarly, leaving out the perfect and building sentences with just the progressive and the passive as in (4), shows exactly the same restriction: there is ‘low’ subject position to the left of the progressive participle.

(4) (a) *There could be being a truck loaded.
 (b) There could be a truck being loaded.
 (c) *There could a truck be being loaded.
 (d) A truck could be being loaded.

The ‘low’ position of the subject is thus at the left edge of a domain that can include the *-ing* participle and the passive participle, but not the perfect participle.

2.1.2 VP fronting and pseudoclefts

Turning now to a distinct phenomenon concerning displacement, it has been argued by Sailor (2012) that VP fronting and specificalional pseudo clefts can target a constituent between Perf *-en* and Prog *-ing* (cf. Sailor 2012). In (5) we see the constituent headed by *-ing* undergoing fronting, and in (6) we see it forming a grammatical cleft. Crucially, the constituent selected by the perfect auxiliary, and that selected by the modal, cannot be targeted in these constructions.

(5) (a) *... [eaten], they will have been being.
 (b) ... [being eaten], they will have been.
 (c) *... [been being eaten], they will have.
 (d) *... [have been being eaten], they will.
 (6) (a) A: John should have been being praised. B: No, ...
 (b) *... [criticized] is what he should have been being.
 (c) ... [being criticized] is what he should have been.
 (d) *... [been being criticized] is what he should have.
 (e) *... [have been being criticized] is what he should.

When the progressive is not present, we see that the constituent consisting of the passive participle can also be fronted much like the progressive participle phrase. Nevertheless, the perfect participle phrase and the infinitival phrase selected by the modal are not legitimate targets.

(7) (a) If Mary says that the cakes will have been eaten, then ...
 (b) ... [eaten], they will have been.
 (c) *... [been eaten], they will have.
 (d) *... [have been eaten], they will.

The examples in (8) show that when both the progressive and passive are present in the absence of the perfect, it is still the *-ing* phrase that fronts. The fact that the passive participle phrase does not front on its own seems to indicate that what is being targeted here is the maximal phrase of a certain type.

(8) (a) If Mary says that the cakes will be being eaten, then ...
 (b) *... [eaten], they will be being.
 (c) ... [being eaten], they will be.
 (d) *... [be being eaten], they will.

These facts show that there is a privileged boundary at the point between Perfect *-en* and Progressive *-ing* which is not dependent on the surface presence of any specific aspectual feature or morphological exponent.

The facts can be modeled by assuming that when they exist, the main verb, passive participle and progressive participle all lie within a particular distinguished domain targeted by these fronting operations. This is the constituent that is fronted in ‘VP -fronting’, and what is clefted in the pseudocleft construction.

2.1.3 British nonfinite *do*-substitution

Finally, I turn to an argument of my own from British nonfinite *do*-substitution, which exposes the same essential division. In British English, *do* is an abstract pro-form that substitutes not just for eventive verbs but for stative verbs as well, after an auxiliary.

(9) (a) John might leave, and Mary might do also.
 (b) John might really like oysters, and Mary might do also.

Although British English *do* can replace stative verbs, it is confined to main verbs and never substitutes for an actual auxiliary. In other words, it is in complementary distribution with stranding by auxiliaries.¹³

(10) (a) John might have seen the movie, and Mary might (**do*) also.
 (b) John might be singing a song, and Mary might (**do*) also.

¹³ Note that the mismatched reading in (10) where *do* is construed as substituting for a main verb in non-finite form after the modal auxiliary, is marginally possible here, but is irrelevant and will be ignored in what follows. The reading where it substitutes for the auxiliary phrase is robustly ungrammatical.

However, even within these constraints, not all nonfinite main verb forms may be substituted for by *do*:

(11) (a) John might leave, and Mary might do also.
 (b) John has left, and Mary has done also.
 (c) John is leaving, and Mary is (*doing) also.
 (d) John was arrested, and Mary was (*done) also.

British nonfinite *do* can substitute for an infinitive modal complement or a perfect participle, but not for a progressive or passive participle. This phenomenon too, motivates a cut between Perf and Prog. The diagnostic is in some sense the converse of the previous one: the very constituents that could participate in the fronting constructions are the ones that British nonfinite *do* cannot substitute for.¹⁴

There is thus robust evidence for two distinct domains from three independent sets of grammatical facts. In each case, the facts point to a joint between the progressive participial phrase and the perfect participial phrase when they exist (and we assume that the joint exists even when the morphological evidence is not so articulated).

Convergent Evidence for the Lowest Zone

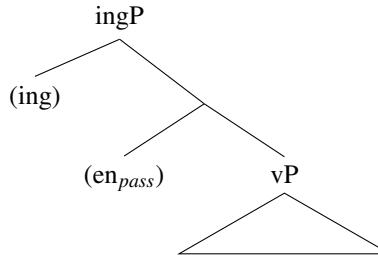
- -*ing*-Phrases, Passive -*en*-phrases and main verb phrases all contain a base position for the external argument
- -*ing*-Phrases, Passive -*en*-phrases and main verb phrases all form a unit with regard to independent mobility
- -*ing*-Phrases, Passive -*en*-phrases and main verb phrases cannot be substituted by the pseudo-auxiliary verb *do* in British English

Thus, with respect to a crude macro division of the clause into a VP-domain and a TP-domain, it seems the progressive and passive forms lie within the lower domain, while modals and the perfect lie within the higher. British English nonfinite *do*-substitution is a pro-form for the higher, but crucially not the lower domain. This makes the difference between the British English dialects and the more restrictive ones, such as the American, quite simple to state: standard dummy *do*-support in the US dialects has only finite instantiations, British English possesses a non-finite version of this *pro*-form as well. If we locate passive -*en* in -*en_{pass}P*, and -*ing* in -*ing_{prog}P*, then the phrase structural

¹⁴ Note that Baltin (2006) shows that British *do*-substitution does pattern like a *pro*-form, rather than ellipsis with respects to the tests in Hankamer and Sag (1976).

description for what we see to see from the purely syntactic evidence given above, can be represented as in (12).

(12) FIRST PHASE:



What we know about selection also supports this division. We have already seen that *akitionsart* is one of the verbal properties that is encoded by lexical items within the verbal domain. As is well known (see e.g. Dowty 1979), the progressive in English selects specifically for the *akitionsart* of its complement—combines with dynamic verbal projections and not stative ones (13).¹⁵

(13) (a) John is dancing the tango.
 (b) *John is knowing the answer.

Under the assumption that selectional restrictions are strictly local (Baltin 1989), the fact that the progressive places selectional restrictions on the *Akitionsart* of the verb phrase it combines with is initial suggestive evidence that *Prog* is low enough in the extended projection to select for the nature of the event structure described by the verb.

However, the phrase structure in (12) is not yet either explanatory or satisfying, because it simply reuses the specific morphological forms as labels and as such is not generalizable to other languages. For example, we want to know whether the projection headed by *en_pass* is actually *Voice*, as described by Kratzer (1996). Similarly, what is the proper abstract label for the projection headed by *-ing*?

I will pursue the natural conjecture, given the proposal in chapter 1, that the lower domain diagnosed here is the domain of abstract eventive properties independent of specific instantiation. This lowest zone denoting properties of

¹⁵ It is equally well known that there are a number of systematic exceptions to this generalizations. One class of ‘exceptions’ is when normally stative verbs are coerced into a more dynamic sense, as in *John is being good today*, or *I am really loving this pizza..* The other class of exceptions concerns the positional statives such as *sit, stand, lie*, which regularly occur in the progressive with non dynamic interpretations. I will not talk about the latter class here.

D_μ contains the progressive and the passive *be*, while the higher domain is the domain of spatio-temporal properties of situations and contains the perfect auxiliary *have*.

Note that the Perfect does not constrain the Aktionsart of its complement. In (14), we see that the perfect can combine with *any* main verb in the English language. While it is true that the meaning of the perfect changes subtly depending on the type of main verb, one could argue that the perfect itself can be technically formed with any main verb regardless of aktionsart, which is different from the progressive.

(14) (a) John has destroyed the castle. (telic verb)
 (b) John has driven on ice (before). (atelic verb)
 (c) John has known Sue for three years. (stative verb)

This goes along with the fact that the perfect has a relationship with temporal anchoring which is different from the progressive, showing a more indirect relationship to the VP event description. In the chapter on the perfect, I will discuss the details of this relationship further. But in brief, an adequate analysis of the perfect is going to require reference to an actual situational particular as a topic situation which mediates the relationship between the event and the utterance event d.

The progressive and the passive are the two lowest in the full possible expansion of auxiliaries in English, and while the passive has traditionally been considered to reside within the first phase zone because of its relationship to the traditional category of Voice, the same has not been systematically claimed for the progressive. It should be pointed out that the progressive shares with the passive the use of the helping auxiliary *be*, which Bjorkmann (2011) argues is inserted as a default to spell out tense features. This makes it more similar to the passive, and less similar to the modals and perfect constructions which introduce their own distinct tense carrying auxiliaries.

Since the analysis I will end up proposing will be a departure from previous approaches, I will begin by describing in the next section the previous semantic treatments of the progressive in the literature. The analysis I will propose should be able to do justice to the issues and subtleties already uncovered by all the foundational work on this topic. However, as I set out to show, it will also avoid some of the bigger problems and paradoxes inherent in the previous approaches.

2.2 The Semantics of the Progressive

As stated earlier, the English progressive is most famous for introducing a puzzle for semanticists called the imperfective paradox. A rich literature has arisen seeking to solve the basic puzzle and account for the increasingly subtle judgements people have about the relationship between a completed event and its progressivized version. The current consensus seems to be that some sort of modalized semantics is necessary. Here are some examples of the denotations proposed in the literature. The first is the classic version proposed in Dowty (1979). He proposes that the progressivized version of the sentence is one which is true if the non-progressivized version *would have been true* in some inertial world that continues on from the current one in the most ‘normal’ way. Crucially, the inertial world in question need not be the actual one.

(15) *Dowty’s (1979) Semantics:*

[Prog ϕ] is true at $< I, w >$ iff for some interval I' s.t. $I \subset I'$ and I is not a final subinterval for I' , and for all w' such that $w' \in \text{Inr}(I, w)$, ϕ is true at $< I', w' >$

where $\text{Inr}(I, w)$ - set of inertial worlds for w and interval I .

Inertia Worlds - are to be thought of as worlds which are exactly like the given world up to the time in question and in which the future course of events after this time develops in ways most compatible with the past course of events.

(from Dowty 1979, pg 148)

The version from Landman (1992) is somewhat more sophisticated and is based on a wider range of tricky examples where people have varying judgements. Landman’s crucial, and correct insight here, is that the ‘inertia’ in question is more event internal than worldly. It turns out that one needs to zero in on the particular event situation and its ingredients when deciding what counts as most ‘normal’, since using the whole world in the calculation of normality, as in Dowty’s semantics, leads to some counterintuitive predictions. For example, when a big truck is hurtling down the street and is about to run over Mary while she is crossing the street, we still allow ourselves to use the progressive even though in the most normal world continuing on from the present moment, she surely does not make it.

(16) *Landman’s (1990) Semantics:*

$[[\text{Prog}(e, P)]]_{w,g} = 1$ iff $\exists f \exists v : < f, v > \in \text{CON}(g(e), w)$ and $[[P]]_{v,g}(f) = 1$.

where $\text{CON}(g(e), w)$ is the continuation branch of $g(e)$ in w .

The crucial notion that does all the work here is that of ‘continuation branch’ for an event in a world w . Once the event stops in the real world, we can ‘continue’ the event by moving over to the next closest world as long as it is ‘reasonable’. If the event stops there as well, we can move over again to the closest world and allow the event to continue. We can cobble together a continuation branch as long as the worlds we hop over to remain ‘close’ and ‘reasonable’, *and* we are still dealing with ‘stages of the same event’. The progressive states that the event in question will indeed culminate if we build one of those continuation branches. See also Varasdi for a discussion of the formal and logical consequences of the move from worlds to events, which makes the latter kind of system into one where the progressive involves the possibility operator, rather than a necessity one.

Still, all of these systems, whether the intensionality is relativized to events or not, involve an operator over possible outcomes. The machinery of possible worlds is invoked to make sense of and to formalize our intuitive grasp of possible event continuations. However, I would argue that the mystery is not dissolved by invoking the possible worlds machinery, we still need to rely on an intuitive understanding of what it means for something to be the ‘stage’ of an event, what counts as ‘reasonable’ world, and how close the world has to be to be ‘close’ enough. In subsequent work, Landman (2008) goes into more detail about what constitutes a ‘stage’ of an event, and is an exposition of the similarities and differences between the tense aspect system of Dutch and English, which is in many ways similar to the agenda of this book. But in the system that Landman applies, expressing the progressive and non progressive as being in some sense parasitic on the ‘same event’ still relies on a notion of ‘cross-temporal identity’ for its definition, even if we put aside the modal component.

“Thus, the notion of cross-temporal identity, as understood here, concerns *what we are willing to regard*, in a context, *as the same event* for the purpose of expressing how often something happened [my italics]. A full axiomatization of the intended notion of cross-temporal identity is beyond the scope of this paper, . . .”
Landman (2008), pg 7

We might feel more comfortable if we go the way advocated in Portner (1998) who proposes to use only the mechanisms and machinery standardly accepted in modal semantics as the primitive relations among worlds with which the intensional semantics of the progressive is defined. In what follows we see a proposal for the progressive meaning which uses the ideas of modal bases and ordering sources to do the job.

(17) *Portner’s (1998) Semantics:*
The modal base: Circ(e) = the set of circumstances relevant to whether

e is completed.

The ordering source: $NI(e)$ = the set of propositions which assert that e does not get interrupted. (No Interruptions)

$Prog(\phi)$ is true at $< i, w >$ iff there is an event e in w s.t. $\tau(e) = i$ and for all worlds $w' \in Best(Circ, NI, e)$, there is an interval i' which includes i as a non-final subinterval such that ϕ is true at $< i', w' >$.

This way, we have some unexplained primitives, but they are the same primitives necessary for the understanding of modal meanings more generally, which we already assume we have. However, to make it work, Portner does have to assume crucially that the construction of the modal base cannot strictly be that shown above, but has to be relativized to the *nature of the event description itself*. Otherwise, the truth conditions of (a) and (b) sentences below would be the same when describing the same event in the world.

(18) (a) Mary was crossing the street.
 (b) Mary was walking into the path of an oncoming truck.

So in fact he argues that, $Circ(e)$ needs to be $Circ(e, P)$, and $Best(Circ, NI, e)$ has to be $Best(Circ, NI, e, P)$. The truth conditions of the progressive therefore come out as (19).

(19) $Prog(e, P)$ is true at a world w iff for all worlds $w' \in Best(Circ, NI, e, P)$, there is an event e' which includes e as a non-final subpart, such that $P(w')(e')$ is true.

While it does seem better to re-use mechanisms we already know to be active in the semantics of other expressions, there are some worrying aspects to the Portner solution that give us a clue that there is something different going on here. The first is the relativization to the *actual event description*, the consequences of which are left at a completely intuitive level.

I think it can be shown in fact that the kinds of calculations that feed into modal judgements are *not* of the same nature as the ones that seem to be involved with the progressive. Specifically, the sensitivity to the nature of the description is quite special here, and the extremely local nature of the worlds calculated over when licensing the progressive. In the street crossing scenario, even though a bus is racing towards Mary, we are still inclined to say that (20-a) is true, but if we avert our eyes at the last moment we would probably say that (20-b) is true, and not that (20-c) is true.

(20) (a) Mary was crossing the street.
 (b) Mary must have been hit by that bus.
 (c) Mary must have crossed the street.

The above argument is due to Klinedinst (2012) who argues against the modal account of the progressive for this reason. Basically, the standards required for asserting the progressive are much lower than the ones we would require to make an epistemic prediction about future events. The calculation we are doing is based in quite detailed ways on the way in which the event itself is presented qua description, and the sense in which the completion is invoked is really as a means of pinning down what *kind* of event is being asserted to be ongoing.

Another aspect of the progressive's semantics that is important to discuss here is the fact, also noted in the literature, that the progressive is 'stative' from the point of view of its external semantic distributional behaviour. As far as I know, the modern treatment that takes this fact most seriously is the one in Hallman (2009a), although he essentially builds on insights from Mittwoch (1988) and Vlach (1981). In Mittwoch's version, the output of progressivization is the construction of a homogenous activity which satisfies the divisibility property.

(21) *Mittwoch's (1988) Semantics:*
 PROG(A) is true in M relative to (w, i) iff i is a subinterval of an interval j and A is true in M relative to (w, j) , where A is interpreted as an activity or state (i.e. homogenous situation).
 from (Mittwoch, 1988, pg 213)

Vlach (1981) is more specific about the stativity of the progressive and includes a processing operator embedded within it.

(22) *Vlach's (1981) Semantics:*
 PROG[ϕ] if and only if STAT[PROC[ϕ] goes on]
 where PROC[ϕ] is that process P that leads to the truth of ϕ
 (Vlach, 1981, pp 287-288)

It is important to note that when it comes to the behaviour of *when*-clauses and other point adverbials, statives pattern one way in English and all the events go the other (including the activities). Therefore, merely divisibility down to a certain grain size is not enough to capture the discourse effects. The progressive specifically patterns like a state, it is not just the kind of homogeneity that we also find in activities. The point about the progressive and *when*-adverbials was originally made by Leech (1971)

(23) (a) When we arrived she made some fresh coffee.
 (b) When we arrived she was making some fresh coffee.

In terms of narrative progression, I show below that the progressive patterns specifically with states and not with activities. In (24-a) and (24-b), the event advances the narrative time in the middle sentence, while in (24-c) and (24-d) the middle sentence does not (cf. also Kamp and Reyle 1993).

(24) (a) John arrived. He sat down. Then he left in a hurry.
 (b) John arrived. He drank coffee. Then he left in a hurry.
 (c) John arrived. He was sweating. Then he left in a hurry.
 (d) John arrived. He looked hot and bothered. Then he left in a hurry.

Hallman (2009a) adds further diagnostics to the stativity claim. For example complements of ECM *discover* and *reveal* in English must be specifically stative and are bad with events of all kinds, including activities. Once again, the progressive patterns with the statives with respect to this test. The following data is from (Hallman, 2009a, pg 8)

(25) (a) The inspector revealed/discovered Max to be a liar.
 (b) The inspector revealed/discovered Max to be lying.
 (c) *The inspector revealed/discovered Max to lie.

In addition, as pointed out by Hallman (2009a), the progressive patterns with statives in being possible in the present tense in English with the same interpretation as the past tense (unlike eventives which shift to a habitual interpretation, or a narrative present). The reason is, as Hallman also argues, statives and the progressives can be true at a ‘point’ in time, while eventives which have duration cannot.

(26) (a) John looked tired when I saw him yesterday and he looks tired now too.
 (b) John was writing a novel when I saw him yesterday, and he is writing a novel now too.
 (c) John ate a mango when I saw him yesterday, and ?? John eats a mango now too.

To cite yet another stativity diagnostic in English, Portner (2003) points out that the universal reading of the perfect is triggered in English for states, and is impossible for events of all stripes including activities. In (27), I use the *since 5 o’clock* phrase to trigger and force the universal reading of the perfect. Only statives and progressives are licit.

(27) (a) John has been in the park since 5 o'clock.
 (b) John has been jogging since 5 o'clock.
 (c) *John has driven a truck since 5 o'clock.
 (d) *John has broken the vase since 5 o'clock.

In Hallman (2009a), he gives an interpretation for the progressive that abstracts away from the possible worlds problem and concentrates only on making sure that the denotation properly expresses that the output of the progressive operator is a state that is only true of points in time (essentially following Taylor 1977). The denotation given in Hallman (2009a) is given in (28).

(28) *Hallman's (2009a) Semantics:*

$$[[\text{PROG}(\phi_{\text{EVENT}})]] = \lambda t \in T \lambda e \in E [\tau(e) = t \wedge \exists e' \in E \exists i \in *T [\Phi(i, e') \wedge \text{Cul}(e, e')]]$$
 where Cul is “intended to be understood as a metavariable for whatever set of circumstances relates the present goings on to the possible culmination, which will inevitably involve reference to possible worlds” (pg 22)

Finally, in separate work, Hallman (2009b) adds a further data issue to the complex problem of how one identifies an event from its processual subpart, by examining data from the interaction with quantification. Hallman notices that the intuitive truth of a progressivized VP depends in an interesting way on the quantificational properties of the direct object and the situation it is claimed to be true of. Consider the sentence below in (29).

(29) The machine rejected one third of the transistors.

Consider this sentence in the scenario where a machine is checking transistors that have been manufactured for flaws. If the transistor is ok it accepts it, and if it finds a mistake it rejects it. In scenario 1, the machine systematically rejects one out of every three transistors it examines in a pattern YES-YES-NO-YES-YES-NO etc. Once the machine has done its job on all 300 transistors, the sentence in (29) would be judged true. In scenario 2, the machine accepts the first 200 and rejects the last 100. Then, too, the sentence in (29) would be judged true. However, consider now the progressivized sentence in (30) uttered after the first 120 trials.

(30) The machine is rejecting one third of the transistors

This can only be true under the first scenario, not under the second *even if* we are watching it on film and we know what the outcome will be.

Hallman's response to this problem is to devise a semantics which utilizes situations and requires of the progressive that the situation of which it is true be divisible (i.e. for the progressive must be true of all its subsituations). (This gives him a problem with accomplishment predicates, which he solves by assuming that the telos for accomplishment predicates is *added* by a null telicizing operator which is in complementary distribution to the English progressivizing operator). The solution he proposes has the result that it predicts that VPs with quantifiers in the scope of the progressive that are not proportional will not be licit. This seems to be true.

(31) (a) ??John was eating exactly three apples.
 (b) ??John was drawing less than ten circles.

Hallman's denotation for the progressive as found in Hallman (2009b) is given below in (32). Note that in this work, he does not directly build in the stative properties of the progressive as argued for in Hallman (2009a).

(32) *Hallman's (2009b) Semantics:*
 $\forall\phi \subseteq [[PROG(\phi)]]^w = \lambda s \leq w \forall s' \leq s R(s', s) \rightsquigarrow \phi(s')$
 where R "essentially represents the 'is a relevant subpart of' relation"

Hallman's analysis of the progressive directly utilizes situations instead of possible worlds (as for example the work on Spanish perfective and imperfective tenses in Cipria and Roberts (2000)). This kind of analysis is helpful in one sense in that it can, like Landman's account, build in the sensitivity to events via the exemplification relation, and does not have to work with all the detail found within possible worlds. Nevertheless Hallman's account still has the property common to the other accounts discussed above that it builds some kind of intensionality into the semantics of the progressive meaning.

Finally, it is worth highlighting the fact that is implicit in a lot of the discussion around progressive meanings so far, which is that native speakers' judgments of the truth of a progressivized sentence are highly dependent on not just the specifics of the description, but also on details of the context of utterance and what kind of information is in the common ground. These kinds of considerations are used by Landman (1992) famously, to argue for the notion of event continuation branch. For example, I am willing to assent to the truth of (33-a) in a particular scenario if I know that Mary is a robot with super-human skills in a science fiction movie where she is sent back in time, but not if she is the girl from the farm next door. If I knew Mary was training hard on her swimming and she just jumped into the water at Dover, I might well agree to (33-b), but not if she is my 6-year old daughter. Also, if I know that Mary was

intending to cross the street, my answer to whether I think (33-c) is true when she sets off would be different from if I know her intention is simply to stand in the middle of the road and block the traffic.

(33) (a) Mary is wiping out the Roman Army.
(b) Mary is swimming the English Channel
(c) Mary is crossing the street.

In a slightly different vein, if I know that digging a deep hole for foundations is the first stage of building a house, I could look at a bunch of workers digging a hole and truly assert “They are building a house.” But people can also disagree in their judgements depending on how well they think a certain activity represents the corresponding named non-progressivized event. For example, I can point to Mary sitting in the library with a huge mound of books on sixteenth century maps and tell you knowingly “Mary is writing an article on maps.” You might disagree with my description of the facts and say. “No she isn’t. I’ll count those publication points when I see them. Let’s hope she eventually does get around to writing, but knowing her I doubt it.”

But one thing remains curiously robust in all of this contextual sensitivity and variability, and that is the fact that people will all agree that the following sentence in (34) is good.

(34) Mary was crossing the bridge when earthquake hit, so she never made it to the other side.

So ways in which the world might be or necessarily must be are irrelevant to our willingness to agree to that statement. Internal facts about Mary, and about her own intentions *are* relevant, but external circumstances are not. This highlights the fact that the judgements here are not equivalent to modal possibility in a general sense, but have to do with our judgement of whether certain ‘essential’ properties of an eventuality are being evidenced or not.

Other well known problems arise with purely modal accounts if the specificity of the event description are not taken into account. For example, the following scenario, attributed to Richmond Thomason in Dowty (1979), is that even though the possible continuations of an event where a coin has just been tossed into the air involve equal numbers of worlds where it comes up heads as when it comes up tails, *neither* of the following two sentences seem to be intuitively true.

(35) (a) The coin is coming up heads.
(b) The coin is coming up tails.

A modal possibility account seems to predict falsely that both should be true.¹⁶ On the other hand, a modal necessity account runs into trouble with the sentences from Abusch (1985) below where the two apparently incompatible outcomes in (36) can *both* be judged true in the progressive in the same situation.

(36) (a) John was crossing the street.
(b) John was walking to his death.

The point is that in (35), neither of those two sentences are good ways of ‘describing’ or identifying a ‘coin toss resolution’ event, given what we know about tossing a fair coin. In (36), on the other hand, *both* of those sentences are legitimate ways of ‘describing’ a ‘walking in front of an oncoming truck’ event, depending on which aspects one wants to emphasize.

In short, we as linguists know a lot more about the difficulties of expressing the meaning of the progressive than we knew before we started on this journey. As a way of summarizing the discussion and paving the way for what is to come, I will here list the core semantic properties/paradoxes to do with progressive meaning that any successful analysis needs to be able to account for.

(37) **Core Semantic Features of the Progressive**

- (i) The progressivized eventuality is related in an organic way to its non-progressivized counterpart, but does not actually entail it (in the actual world) at a future time. Judgements of event sameness are due to some judgement of ‘essential identity’, rather than prediction of outcome.
- (ii) The relationship between a progressivized event and the event simpliciter is not qualitatively the same as epistemic uncertainty (Klinedinst’s Observation).
- (iii) The perceived relationship between a progressivized event and the event simpliciter is affected by contextual properties of the discourse and gives rise to variable judgements across speakers. In this regard, *internal properties of the participants and their intentions*, and the *nature of the process evidenced* seem to be more important than external circumstances. (Landman’s Observation)

¹⁶ This is essentially a version of the ‘multiple choice paradox’ as later discussed and explored by Bonomi (1997): the progressive of a disjunctive set of options can be true in contexts where none of the progressivized versions of the individual options would be judged true.

(iv)The progressive functions like a state in its temporal semantics
(Vlach, Parsons, Hallman)

I think it is fair to say that all of the possible worlds accounts we have seen fall short of complete objective explicitness when it comes to point (i) above. In all cases, the appeal to possible worlds still leaves an unexplained residue completely independent of the possible worlds mechanisms themselves. In the case of Landman it is his appeal to the ‘stage-of’ relation, in Portner it is the relativization to event descriptions, in Hallman’s situational version it is the relation R ‘the relevant subpart relation’. The essential question of “What does it mean to be an in-progress version of an event?” remains a primitive, and essentially a mystery.

The possible worlds framework of Lewis and most subsequent formal semanticists has appealed to the field because of its formal explicitness. Possible worlds allow us to extend the ordinary notations for extensional reality and express hypotheticals and uncertainties by means of regularities over multiplicities of worlds. The core of the semantics remains extensional, it is just the worlds that are hypothetical/possible/likely etc. Possible worlds are a convenient fiction which seem on the face of it to make the formalization of many aspects of natural language more tractable, elegant and well suited to compositional treatment. But in the case of the progressive, the possible worlds accounts are much less compelling. A basic mystery remains at the heart of all of these accounts, which I would argue is this judgement we humans make instinctively regarding essential event identity.

In my own proposal, I will assume the equivalent of the unexplained part as a basic cognitive primitive. To anticipate, the ability to identify a snapshot state of an event as being a *part* of that event, is a sensory/cognitive judgement that forms the basis of our ability to classify the world based on symbolic labels. In an important way, the proposal I will make is a close sister of the position taken in Parsons (1990), although I will define my equivalent of Parsons’ ‘in-progress-state’ somewhat differently from him, avoiding the mereological part-of relation.

Taking a particular cognitive judgement to underlie a primitive property does not necessarily make the analysis any more objectively verifiable or accurate, although I do think it makes predictions as well as other accounts. It will have a number of positive payoffs with respect to polysemy and the interaction of the *-ing* participle with the rest of the grammatical system. It will also be the starting point for testable predictions within cognitive psychology and acquisition .

2.3 The Present Proposal

Before giving my own account, I will present the account of Parsons (1990) and the objections to it that have been raised in the literature. The Parsons account is intuitively speaking the closest intellectual precursor to the present one. I will discuss the criticisms of that account, and use them to define my own version that avoids what I see as the major problems.

Parsons (1990) avoids modality in his account of the progressive by assuming the notion of incomplete event as a primitive. Actually, technically speaking, he assumes the predicates **Hold** and **Cul** as primitives and asserts that ‘incomplete’ events and objects still count as those very events and objects as far as speakers are concerned. In (38-a) and (38-b), we see the representations Parsons proposes for the progressive sentence and the ordinary past tense accomplishment sentences respectively.

(38) *Parsons 1990*

(a) Mary was crossing the street.

$\exists e \exists I \exists t [I < \text{now} \wedge \text{cross}(e) \wedge \text{Agent}(e) = \text{Mary} \wedge \text{Theme}(e) = \text{the-street} \wedge \text{Hold}(e,t)]$

(b) Mary crossed the street.

$\exists e \exists I \exists t [I < \text{now} \wedge \text{cross}(e) \wedge \text{Agent}(e) = \text{Mary} \wedge \text{Theme}(e) = \text{the-street} \wedge \text{Cul}(e,t)]$

The Parsons account avoids the imperfective paradox because (38-a) does not entail (38-b). In the semantics of **Hold**, there is no entailment that **Cul** ever becomes true.

It has been objected that this account simply has nothing to say about the actual judgements people have about (38-a), namely *when* it is they would agree to its being true. I submit that this charge is unfair. Parsons would say that people judge (38-a) to be true precisely in the circumstances in which they would agree that (38-a) is a *part* of an event that could be described by (38-b). We have seen that all the modalized accounts reduce to the equivalent of this mysterious statement as well.

Another objection to this account in the literature concerns our intuitions about incomplete objects. The idea is that from the denotations given it seems as if Parsons is committed to the idea that if ‘John was drawing a circle’ , then there is ‘a circle’ that he was drawing, even if he has just begun to draw. Parsons bites the bullet here and says that in some sense this is true, there *is* a circle, albeit an incomplete one, just as there is an incomplete event of drawing a circle. Many people feel uncomfortable with this judgement.

Another version of the objection also zeroes in on the problematic aspects of the notion of partiality, but with respect to the event denotation. The objection is due to Zucchi (1999), who points out that the Parsons account makes a prediction about the underlying denotation of the bare verb phrase, namely that it should be able to denote partial events, at least optionally. This is because Parsons assumes an underspecified semantics for the VP and then allows the predicates **Hold** and **Cul** to build unambiguously atelic vs. telic events respectively. On the other hand, the modal accounts assume a completed event denotation and use the progressive operator to remove the entailment to completion in the real world. To the extent that bare untensed verb forms can be found in English, the data seem to favour the modalized account. In other words, the bare VPs in (39) seem to be interpreted as ‘complete’ events.

(39) (a) John saw Mary cross the street. # He saw the bus hitting her when she was halfway across.
 (b) John saw Mary crossing the street. He saw the bus hitting her when she was halfway across.

So is this the final nail in the coffin of the Parsons account? I think it speaks against the particular implementation that Parsons offers in terms of partiality, but it will not apply to the version I will describe next in terms of linguistic event concepts and Identifying-States.

2.3.1 The Progressive as an ‘Identifying State’

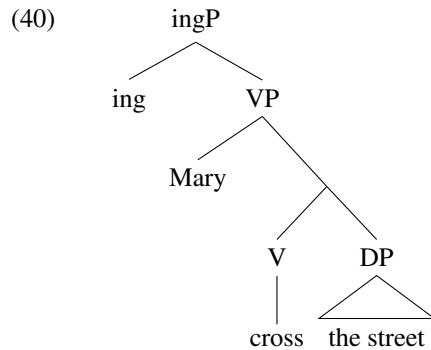
The Identifying-States account is going to share with Parsons (1990) the idea that the relationship between the progressivized version of an event description and the underlying event description is not to be described in terms of possible worlds.

First of all, as I have argued in chapter 1, we use the lowest phase of the syntactic representation—the first phase—as the level at which linguistic items are combined to create a symbolic constraining description of the event being conveyed. The information expressed by our verbal linguistic items at this level is devoid of temporal or situated content, but just represents a classification at the level of basic sensory/perceptual commonalities. However, I think the empirical evidence shows that the domain of linguistic symbols D_μ is complex, hierarchically structured and is a part of a generative, compositional system. The question is what sorts of compositional relationships exist at this level.

We need therefore, to be explicit about the relationship between the *-ing* form of a verb, and the tensed form of the verb. The *-ing* form of the verb is the crucial linguistic ingredient in the formation of what is known as the English progressive. The relationship between the progressivized event description and

the bare event description cannot be in terms of **Hold** and **Cul** as in Parsons treatment, because these are notions that apply to instantiated, temporal entities and not between atemporal event descriptions.

The task therefore is to express the denotation of the *-ing* participle, as an *ingredient* of the progressive. I will assume first of all that the *-ing* morphology is a productive suffix that applies and takes scope over the whole VP constructed so far with all of its required arguments.



Taking seriously evidence for the stative nature of the progressive in English, I propose that the semantic part of the *-ing* morpheme is a function from event descriptions to event descriptions such that the derived eventuality is an **Identifying-State** for that event.

In the description of the ‘quotational’ semantics in chapter 1, linguistic items are elements of the ontology and can be composed to give complex linguistic items with a derived conceptual contribution. By default, I have assumed that the standard way of combining the semantic content of the individual pieces is by simple argument identification (of the event argument variable). In the case of *ing*, I will assume a slightly special rule for its composition with a complex phrase within the D_μ domain.

Box 2.1
The Formation of the -ing Participle

(41) If A is formed from the merge of *ing* and B where $B \in D_\mu$, then A is also in D_μ , and
 $\llcorner A \lrcorner = \lambda e[\text{State}(e) \wedge e \text{ is an Identifying-State for property } \llcorner B \lrcorner]$

The definition of an **Identifying State** is given in (42).

Box 2.2**Identifying State (ID-STATE): Definition**

For all event descriptions P , an *Identifying-State* for P , is a stative eventuality that manifests sufficient cognitive/perceptual identifiers of the event property P .

Thus, the denotation of an *-ing*-participle would be as shown in (42).

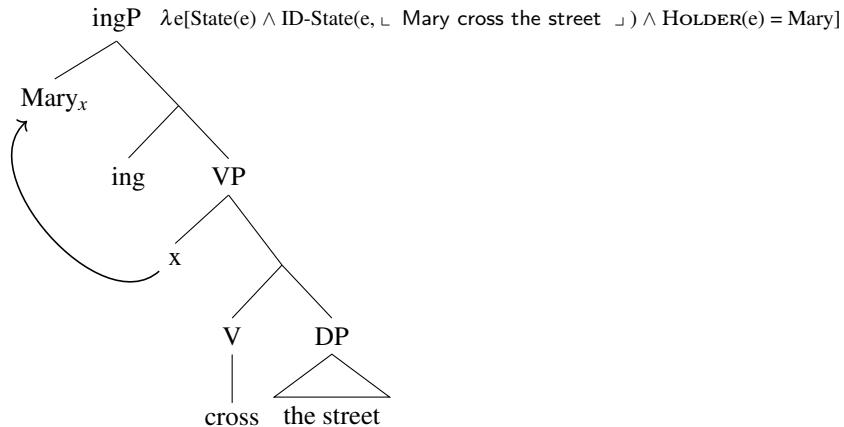
$$(42) \quad \llcorner \ u_V\text{-}ing \ \llcorner = \lambda e[\text{State}(e) \wedge \text{ID-State}(e, \llcorner \ u_V \ \llcorner)]$$

Intuitively, the Identifying-State relation is one in which speakers can infer the existence of a (possibly complex) event E by means of a state that provides sufficient evidence (given everything else we know about the world) for that type of event. These states are conceptual snapshots, which, if rich enough or with rich enough insider knowledge, license inferences to the existence of a more complex extended eventuality in practice, given the right conditions. I assume that ID-STATE is a primitive relation, intuitively accessible to the youngest of children, and one that is at work every time we make claims about durational eventualities based on immediate perceptual contact.

In addition to the shift to states, we need to shift the argument structure so that there is a designated ‘holder’ of the derived state. This, empirically speaking, is always the ‘highest’ argument in the already built-up event description. I assume that the merge of *-ing* triggers the movement of an argument to its specifier, which is then interpreted as the HOLDER of the Identifying-State.¹⁷

$$(43) \quad \llcorner \ u_V\text{-}ing \ \llcorner = \lambda x \lambda e[\text{State}(e) \wedge \text{ID-State}(e, \llcorner \ u_V \ \llcorner) \wedge \text{HOLDER}(e)=x]$$

¹⁷ Following Ramchand (2008), I will assume that DPs that A-move from one argumental position to another are considered to be identical in referent and simply accrue argument entailments additively.



It is easy to see that this kind of analysis does not give rise to the imperfective paradox since the *-ing* event bears a stative identifying relationship to the non-*ing* event property, but neither of them so far is asserted to exist in any particular world or time, so no entailments automatically exist between the instantiated versions.

The notion of ‘Identifying-State’ as defined above also makes explicit that what is at stake in the variability of the judgements we have seen so far is the flexible and slippery nature of ‘evidence’, which is affected by our knowledge of the intentions, proclivities and abilities of the participants as well as basic common ground facts about how event types proceed in the world normally or prototypically. Because the judgement is about the relationship between an event property *P* to its evidential/identifying state, facts about the world that would affect full temporal instantiation of the event property are irrelevant.

Thus, the account satisfies all of the four desired properties I outlined in (37). Moreover, it also gives us a handle on Hallman’s Quantification Problem. This is because of the problem of evidence. If Mary is eating two apples, the evidence that two apples are involved is entirely lacking in a completely serialized eating scenario, but visible and present in a situation where she bites them both simultaneously, or where we see them together on her plate, or when we know her intentions. These facts, I think conform to our intuitive judgements. The divisibility of the enclosing situational description (up to a certain granularity) is a kind of guarantee of the persistence of evidence.

2.3.2 Relating Event Properties to Situational Particulars

Once we sever event descriptions from time and worldly instantiation, we are in an entailment vacuum. Judgements of entailments from one propositional form of English to another are simply not predicted. We have exploited this and used it as a virtue in getting rid of the unwanted entailments from the progressivized version of the sentence to the simple past, but we are in danger of predicting no relationship at all between *-ing* forms and tensed forms unless we spell out the relationship between the existentially bound and anchored *-ing* forms and the corresponding existentially bound and anchored non-*ing* forms. What we know is that that this relationship is sensitive to the *aktskala* of the base non-*ing* form. Sometimes the two forms are related by entailment and sometimes they are not. How do we make sense of this?

The first core fact is that a past tense utterance of any eventive verb in English will entail the past tense version of its progressivized counterpart. Thus, (44-a) entails (44-b) for any (non-stative) verb of English.

- (44) (a) John built a house.
- (b) John was building a house.

I will assume the following natural relationship between events that fully instantiate a particular property and identifying states for that property.

- (45) **Event Existence Entails Existence of Identifying State:**

The existence of an event entails the existence of at least one identifying state. The state in question is always a mereological subpart of that event.

Importantly, the converse does not hold. The existence of an identifying state for an event property does not guarantee the existence of an event that fully exemplifies the property. This means in particular that (45-b) will not in general entail (45-a), although (45-a) entails (45-b).

On the other hand, it is usually assumed to be the case that (46-a) does entail (46-b).

- (46) (a) John was running.
- (b) John ran.

We also need to be able to explain the judgements for activities here, and why they differ from achievements and accomplishments. In fact, I would like to propose that (46-a) does not actually entail (46-b) at all. It is just that the fact that the progressive does not entail the simple past version is much more obvious in the case of accomplishments than in the case of activities. We can

explain this effect because of *inferences* based on real world information combined with the homogenous properties of activities—the fact that if they are true at any interval at all, they are true at *every* subinterval of that interval larger than a moment, including extremely short intervals indeed. The following set of meaning postulates for different primitive *aktionsart* categories of events are given below, adapted from Taylor (1977). They constrain the possibilities for temporal instantiation for the different eventuality types.

Box 2.3**Temporal Properties of Different Primitive Event Types****(I). Temporal Properties of Simple Dynamic Events:**

A process event must have a temporal parameter longer than a moment. If a simple process is true at an interval I, then it is true at every subinterval of that interval larger than a moment.

(II). Temporal Properties of States:

A state can have a moment as its temporal parameter. If a state is true at an interval I, then it is true at every subinterval of that interval, including at each moment.

(III) Temporal Properties of Complex Events:

An event with complex subevental structure must have temporal run times corresponding to *each* of the subevents in that structure. If a complex event is true at an interval I, then we cannot guarantee that there is any subinterval of I at which the complex event is true.

Consider now how the property of activities interacts with the definition of identifying state assumed above. If (46-a) is true then it means that an identifying state exists. Minimally, then it could hold for exactly one asserted time moment. The existential closure of this event variable does not mean that it is unique. In fact, if there is even one other such temporally abutting identifying state for the event property, then we could form the join of those two situations to form something that *would* be an instance of the event itself. Thus the non-progressivized version would be true. In context, the inference is that the real world situation that would support an utterance of (46-a) will also support an utterance of (46-a) one second later. Because of the homogeneity of activities, that inference is enough to guarantee the move to (46-b) in the case of activities which are true at every subinterval larger than a moment, and depending on the context it may even be enough to guarantee such inferences for activities with larger minimal subparts such as *dance a waltz*. But in this story, there is no actual *entailment* between (46-a) to (46-b) just a strong inference, which in certain cases can be denied. If the music started up in waltz-time and John went out intentionally on the floor with Mary and took the first step, then I could see that and utter truthfully *Look, John is waltzing!*. But if he crumples to the floor with cramp after that first step, then we are unlikely to agree after the fact that he actually ‘waltzed’. This is what we predict.

Finally, I need to say a few words about the status of the Object domain in this kind of account which treats the lowest part of the verbal extended projection as compositions of linguistic items. In the semantic part of the linguistic item, we find by hypothesis the specification of participant relations that relate entities from the regular Object domain D_e to events. We need to ask how nominal projections compose with verbal projections to build descriptions of full events. In fact, the rethink I am proposing requires a parallel rethink in the domain of object entities. At the lowest part of the nominal extended projection, we need to be trafficking in nominal linguistic items introducing conceptual content as abstractions over actual instances of referents. It is plausible that in the first phase of the verbal extended projection, we are dealing only with entities of type D_μ . Instead of instantiated objects as participants (in domain D). Movement to the second, inflectional, domain of the clause would then be required to anchor this content to actual claims about particular referents. This assumption will avoid the problem that the Parsons (1990) account faced with ‘incomplete’ objects, because no ‘actual’ object will be asserted to exist, once we make this move. Just as event concepts get ‘instantiated’ as particulars only in the inflectional domain of the clause, so too do nominal extended projections only get actualized at the higher levels of the functional structure, and then related to the situational domain by abstract case licensing.¹⁸

Much recent work on the syntax-semantics interface of nominal projections has indeed proposed a zoning of the nominal extended projection in much the same way as I will be proposing for the verbal extended projection, an idea already present in Grimshaw (1990). In particular, see Zamparelli (2000), Borer (2005), Pfaff (2015b) for arguments that the functional sequence of the nominal projection contains zones for kind level meanings or substances, followed by packaging and counting mechanisms. As with verbal genericity, there has been a debate in the literature about whether nominal genericity should be handled quantificationally or whether there is a notion corresponding to the primitive ‘kind’. Following Fine, I am taking the position that, consistently, there is a level of ‘deep genericity’ within (both the verbal and) the nominal domain that really involves a kind of essential non-instantiated description, and that this kind of deep generic interpretation, or ‘kind’ interpretation if you will, is distinct from the higher more complex forms of ‘quantificational genericity’.

¹⁸ The details of the interleaving of the extended projection of DP and VP are beyond the scope of this monograph however. I assume that even though the nominal projection is encapsulated away from the extended verbal projection, the existence of higher levels of functional structure which are not interpreted at the lower levels is what drives movement to the higher domains of the verbal functional sequence.

It is captured by the fact that at the lowest level, we are trafficking in partial nominal properties and properties of actual referents.

In the next section I show how the denotation given for *-ing* above can carry over plausibly to the gerundive uses and nominalized uses of the *-ing* forms in English. This represents a simplification of the grammar that previous accounts do not attempt. It is a unification made possible by the association of the *ing* morpheme in English with the idea of Identifying States.

So far, the proposal is that the morpheme *-ing* in English is a function from D_μ to D_μ which has the following denotation:

$$(47) \quad \llcorner \ u_V\text{-}ing \llcorner = \lambda x \lambda e [\text{State}(e) \wedge \text{ID-State}(e, \llcorner \ u_V \llcorner) \wedge \text{HOLDER}(e) = x]$$

where the value for x is filled in by the movement of the highest argument to the specifier of the *-ingP*.

Notice that there are some well known things about the progressive that I have *not* built into this denotation. One is the restriction to dynamic events that the English progressive is famous for. An obvious way to build in this selectional requirement would just be as a brute force presupposition on the nature of the P that *-ing* combines with. However, there is compelling evidence that the restriction to dynamic eventualities is a property of the progressive construction as a whole, not of the *-ing* participle per se. Even if we look at the most closely related form, the reduced relative or attributive participle in *-ing*, we see immediately that there is no ban on stative verb phrases as the input to *-ing*.

(48) (a) A man is dancing in the corner.
 (b) A man is eating an apple.
 (c) *The wall is surrounding the castle.
 (d) *The boy is fearing the dark.

(49) (a) The man dancing in the corner is tall.
 (b) The man eating an apple is tall.
 (c) The wall surrounding the castle is high.
 (d) The boy fearing the dark was the only one who could not get to sleep.

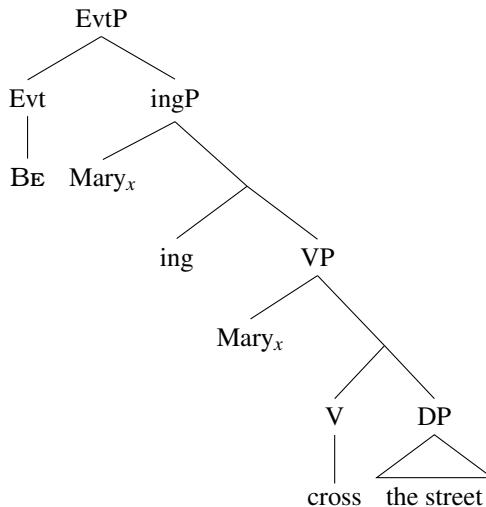
For this reason, I will not build a presupposition into the denotation of *-ing* itself. Instead, I will assume that the ban on (49-c) and (49-d) is due to the

competition for lexical insertion with the simple tensed forms in the case of stative verbs.¹⁹

The progressive is composed of the participle in *-ing* and the progressive auxiliary *be*. I will essentially follow the analysis of Bjorkmann (2011) in proposing that the *be* auxiliary is inserted as a dummy verb in order to host inflectional features whose exponence is required. Participle formation in *-ing* does more than create a derived state-description, it also has syntactic effects. Specifically, I will assume that the output of merge with *-ing* no longer has unvalued features for tense or aspect.²⁰ On the assumption that tense information needs to be expressed in English for sentential well-formedness, a dummy verb must be inserted.²¹

Recall from chapter 1, we introduced an Evt head at the edge of the D_μ domain which introduces the utterance situation d and which expresses the relation of deployment of u in conveying e .

(50)



¹⁹ Essentially an extension of the phenomenon of ‘Poser blocking’ (Poser 1992). See also Lundquist (2008) for a discussion of the formation of the equivalent of *-ing* participles in Swedish, and their various polysemies, including a discussion of blocking. For Lundquist, the Swedish equivalent of *-ing* is simply a nominal gender feature that can attach to structures of various sizes.

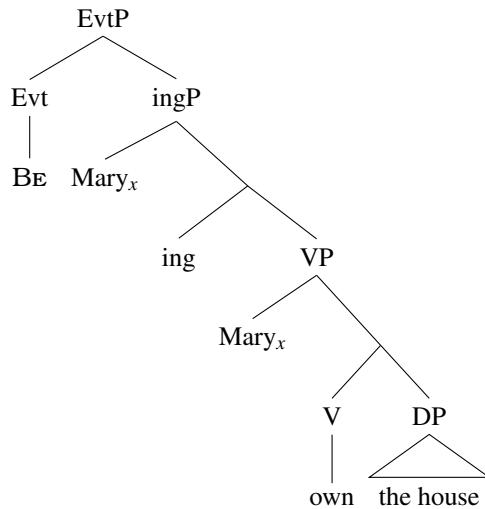
²⁰ I leave technical details of implementation aside here.

²¹ One could assume further that the inflectional information present on main verbs in English is placed there by downward Agree from higher interpretable heads. Another alternative would be to say that the main verb does indeed bear T, but that the spell-out diacritic for the T-Asp-Evt-etc. span in English is always located at Asp.

We can now account for the distribution of auxiliary *be* by saying that it spells out Evt in the absence of any other overt exponent. *Be* contains unvalued features and can host both aspectual and tense information. *Be* otherwise has no semantics.

If we were to build the same phrase structure for a stative verb, we would build (51-a) with the semantics in (51-b).²²

(51) (a)



(b) $\lambda e[\text{State}(e) \wedge \text{ID-State}(e, \perp \text{ Mary own the house } \perp) \wedge \text{HOLDER}(e) = \text{Mary}]$

Since we have assumed that a definitional property of states is that they are true at a moment, the existence of the ID-state for a property entails the existence of the state, and vice versa. Thus, for states and states alone the two notions are mutually entailing. We can state this explicitly as an axiom of this system, although I suspect it can be derived by the fact that the lexical stative property is adduced by the very same sort of evidence that the ID-state requires, so that there could be no difference between the two. This is expressed in (52).

(52) If $u_V \in D_\mu$ is a state semantically, then
 $\perp u_V \perp = \perp u_V\text{-ing } \perp$
 (i.e. The ID-State of s is just s)

²² The objects of stative verbs are part of the event description of the state. They are in what Ramchand (2008) calls ‘rhematic’ position.

It is only in the case of states that there is no difference between the state itself and the identifying state, with more complex events, the identifying and inferential properties of the ID-state are distinct from the conceptual properties of the whole extended eventuality it is related to (the latter expressing unfolding properties in time, while the former is static).

This means that the semantics of the tree in (52) is strictly identical to the semantics that would be derived by (53).

(53) Mary owns the house.
 $\lambda e[\text{own-the-house}(e) \wedge \text{HOLDER}(e)=\text{Mary}]$

I will assume that this means that by some reasonable statement of semantic economy, that the attachment of *-ing* is prohibited in precisely this case.

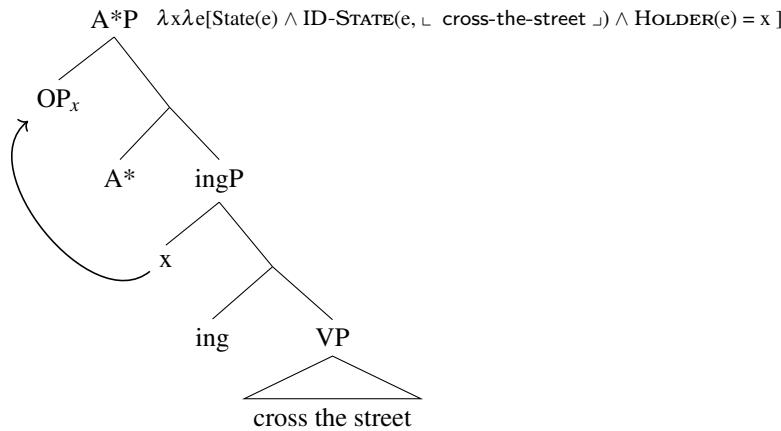
When we look at Passive formation in the next chapter, we will see that the passive auxiliary *be* is the very same lexical item *be* that is inserted because of the participle formation in the progressive. Thus, it looks like it is the form that is inserted not just here, but in any head within the first phase that fails to have its own overt exponent. The insertion of the dummy verb *be* is therefore sensitive to the semantic zone of the clause. In particular, *be* will be inserted in the first phase, while *do* will be inserted for unfilled heads in the situational domain proper.

The simple denotation for *-ing* given in (41) as an ingredient of the progressive can then be seen as the same input to attributive participle formation provided that we allow it to be embedded under an abstraction operator, which I will call A^* (to invoke the idea of adjectivalization).

The abstraction operator A^* will abstract over the highest argument position in the event description it combines with to create a derived property of individuals. One of the core things about the *-ing* participle in English is that it always modifies the argument that would end up as the subject of the corresponding active verb.²³ I capture this fact with two properties of the *-ing* lexical item. The first is that it combines with an already complete event description (i.e. one that already has all of the arguments present). This is simply a matter of height of attachment. The second property is the fact that it can be selected by the property abstraction operator. I assume that nothing need be stipulated about A^* except for the semantics of predicate abstraction, and the restriction to states. In the denotation that follows, P stands in for the event description contributed by the non-progessivized VP. The crucial point is that it is always

²³ This is in contrast to the *-ed* participle, which we will look at in the next chapter, which always abstracts over the internal argument of the related active verb when used attributively.

the **HOLDER** of the progressivized event description that gets abstracted over in attributive participle formation in *-ing*— the position abstracted over is thus always the most local one to the A* head.



Finally, the participle in *-ing* is also found in nominalized uses. Here again, as expected, the ban against stative inputs is lifted, and the form is found in gerundive phrases with the external distribution of nominalized clauses. As in the attributive use, the external argument position is often missing, but not always (see (54-a)). (I assume this will fall out from independent facts concerning Case/DP licensing in non-tensed and nominalized environments, and will not make any proposals about the choice here.)

(54) (a) John eating all the chocolates was a good thing.
 (b) John's eating all the chocolates was a good thing.
 (c) Eating chocolate is good for you.
 (d) Eating chocolate all the way, I eventually reached the top of the mountain.
 (e) The eating of chocolates is considered to be good for the health.
 (f) John's eating of the chocolates was quick and messy.

This is not the place to discuss the syntax and semantics of nominalization. It would take us too far afield, and I do not have anything to add about the combinatorics that goes beyond the basic intuition in Abney (1987), but see

Borer (2013) for a more recent discussion of the properties of nominalizing *-ing*.

I do however wish to point out that the proposal for the *-ing* denotation given above is suitable for a range of nominalized uses because of its identity semantics. Moreover, the underspecified denotation I have given it does not require any possible worlds, or event partiality, but is defined in terms of the notion of **Identifying-State**. This means that the *-ing* participle bears an *identifying* relationship to the corresponding event, and can therefore be used to name it. The *-en/ed* participle does not have nominalizing uses in English, although it does have adjectival, attributive ones. In addition, as we will see in the next chapter, the semantics for the *-en/ed* is highly underspecified and does not involve any ‘identity’-semantics. I therefore conclude tentatively that the notion of identification is at the heart of the meaning for *-ing* and that this makes it suitable for both nominalizing and progressive participial functions. Since the *-en/ed* participle does not have nominalized uses, I will not derive the nominalized versions of *-ing* by general mechanisms from participles. But I remain cautiously optimistic that the semantics of ID-STATE can be then further input to a generalized nominalizing operation in English.

2.4 Conclusion

In summary, this chapter has argued that we get a better account of the participle in *-ing* if we abandon explicitly intensional accounts and couch the analysis instead in terms of event concepts contributed via the deployment of lexical items that express partial event descriptions independent of time and place and by the construction of derived identifying states.

An important point that speaks directly in favour of this kind of account is that fact that it has been known for a long time that the progressive participle in *-ing* is one of the very earliest pieces of morphology acquired by English children. It is acquired between the ages of 19 - 28 months, and appears *before* both irregular past tense (which in turn appears often before regular past inflection) and the copula (Brown 1973; Owens 2001). The use of the *-ing* participle thus appears before any actual tense inflection or modal expression, and is used correctly immediately.²⁴ A fully modal and intensionalized analysis of the progressive would require us to believe that English children acquire a modalized meaning accurately before they are two years old, and always do so before they even have the ability to express tense or use modal auxil-

²⁴ When it appears it appears first without the helping *be* verb, and it seems to occur first in telic verbs and then is gradually extended to verbs without salient endpoints. It is never apparently overextended to stative verbs.

iaries. The pragmatic complexity of inferences connected to the setting up of modal bases and ordering sources is supposed to be something that children need some social and interactional maturity in order to develop. But standard accounts seem to assume that they can do this even before they pass theory of mind tests! The role of *-ing* in identifying and naming complex generalizations over events seems like a good candidate for the type of meanings necessary at the very earliest stages of language learning.

There are also advantages of the present proposal in that we can get a unified denotation for the *-ing* participle as is it used in the progressive, the attributive participle and a range of gerundive constructions. In each case the participle is embedded within enclosing syntactic structure which adds to the distributional properties of the construction. In the case of the progressive, what is added is a dummy verb which is inserted to carry tense information and the structure is built into a full proposition. In the case of the attributive participle, the *-ing* phrase is embedded under the abstraction operator A^* . In the case of the various gerundive forms, my assumption is that we can use the very same *-ing* phrase we build here as well, but embedded under an independent nominalizing head N^* (also assumed to be null here). The hope is that the different kinds of nominalization and gerundive formed with *-ing* will fall out from the height at which the nominalization is attached. Given that the nominal functional sequence is intended to interleave with the verbal one, and given that nominal structure can extend into the second phase, there are a number of different possibilities that arise. However, this is clearly beyond the scope of the present monograph. However, the point about the denotation of *-ing* is that, as an identifying state for the corresponding event property, it can be seen as an ingredient for these forms in addition to the verbal ones explicitly treated in this monograph.

We have seen that the syntactic evidence from substitution, movement and clefting tell us that the *-ing* participle lies deep within the clause, before the first phase is complete. An account of the semantics of the progressive in terms of possible worlds would not actually dissolve the mystery of the imperfective paradox, and at the same time forces us to build in modal meanings at the very lowest domain of the sentence. Typological evidence from verbal morphology and inflection however show us that modal and tense morphology reside high in the clause. The English progressive is low, and is the first morpheme acquired by children, already at the two-word stage, and before any other tense or modal inflection.

All the evidence converges on a simple primitive relationship between a verbal root and its productive *-ing* derivative. I have argued that the best way to think about this relationship is in terms of the notion of 'identifying states'.

3 The Passive and its Participle

The next two chapters deal with the auxiliary constructions in English that employ the participle in *-en/ed*, sometimes called the ‘passive participle’ and sometimes called the ‘perfect participle’ depending on whether it appears in the passive or perfect construction respectively (1-a) and (1-b). The participle in question is the same lexical item in English in that it is formally identical in both these uses with no allomorphic variation, as indeed it is in the attributive use as well (1-d).

- (1) (a) The cat was chased by the dog.
- (b) The dog has chased the cat.
- (c) The cat has been chased by the dog.
- (d) The recently chased cat....

Much recent work on the passive participle has focused on the difference between the ‘adjectival’ passive and the ‘verbal’ passive (see Wasow 1977), and more recently on the difference between stative target state passives and stative resultant state passives (Kratzer 2000, Embick 2004). While early work explored the idea that adjectival passives are formed ‘in the lexicon’ while verbal passives are built up constructionally in the syntax, more recent research has cast doubt on this idea as a way of explaining the different semantic and distributional properties of this form (Bruening 2014). Instead, in constructivist approaches such as Distributed Morphology, it has become standard to analyse the difference in terms of different heights of attachment of the participial morphology in the syntactic tree (Embick 2004, Anagnostopoulou 2003). Whether one attempts to express this syntactically or lexically, there are a number of acknowledged differences of usage that need to be accounted for. The challenge lies in accounting for the differences while still maintaining a common abstract core representation. To my knowledge, the assumption of an underlying unity for this participle has not so far extended to the use of

the participle in the perfect.²⁵ However, the formal identity of the two participles is at least suggestive, and in this chapter I will propose an analysis that strives for a unified denotation, locating the differences in behaviour in the structural size of the participle. I acknowledge here previous unpublished ideas by Michal Starke (CASTL seminars) who early championed the idea to us that the EN/ED participle in English, up to and including the past tense form was actually the very same item spelling out successively different sizes of structure. This indeed is the view that I will build up to over the course of the next two chapters. However, since I have been unable to establish any concrete proposal for the syntactic labels or denotations of all the heads involved in such spans, I have been unable to compare my implementation of that intuition with Starke's specific proposal.²⁶ The structure of this chapter is as follows. In section 3.1, I will examine the uses of the *en/ed*-participle in stative predicative uses. In section 3.2, I will give an analysis of the eventive passive construction within the first phase, giving an interpretation of the participle that is unified with the stative version. I will also give analyses of the role of the ingredients such as the *be* auxiliary and the Evt head, which are the same as we saw in the progressive chapter. Section 3.3 treats the use of the passive participle as input to adjectivalization.

In the chapter that immediately follows this one, I will attempt to analyse the English perfect, with the very same participle as ingredient. In general, I will seek a unified representation of the *-en/ed* participle, which is why the disparate constructions of the 'passive' and the 'perfect' are treated one after the other. The strategy will be to keep the meaning of the participle constant and use different heights of attachment and the different semantic properties of the zones to ground the known semantic differences. In addition, I will argue that *have* is merged in the situational domain, which in turn will have important consequences for temporal modification and interpretation.

But before we get to the domain of situations, we must discuss the passive participle in some detail, which lies as I will argue, in the lowest domain of D_μ , event concepts.

²⁵ In Swedish, a closely related Germanic language, the same form is *not* used for both passive and perfect. Swedish has a dedicated participial form, 'the supine' which is used in the perfect construction, showing at least that the lexical choice for this kind of participle *can* be different from the passive. For this reason, it has perhaps seemed less urgent to bake in a formal unity between the two forms in a language like English.

²⁶ Starke's own framework involves spell out of constituents instead of spans, which makes it very difficult to construct a system that will integrate the participle with the auxiliaries, especially cases of multiple auxiliaries. For this reason, Starke had to abandon the cleanest and simplest implementation of his intuition (pc class seminars). The present account will use spans which allows a more direct version of Starke's earliest intuition, which I believe to be the correct one.

3.1 The Stative Participle in *-en/ed*

One crucial semantic distinction that has been noted between different uses of the ‘passive’ participle²⁷ in predicative position is whether the resulting predication is stative or dynamic. I will start with examining the stative participle that can be formed with *-en/ed* in English. The eventive passive will be treated in section 3.2.

A complication that we face immediately is that in the literature, a number of different types of stative participle have been claimed to exist. In Embick (2004), these are called ‘resultative’ vs. ‘stative’ participles (although they are both actually stative, as he acknowledges). In Kratzer (2000) a ‘resultant state’ vs. a ‘target state’ passive participle are distinguished. In fact, these two authors are not making precisely the same distinction with these labels, so we cannot simply choose our terminology here. In the case of Embick, what is important in distinguishing the two classes is the presence or absence of ‘event implications’: resultative stative participles have event implications, ‘pure stative’ participles do not. In the case of Kratzer, what is important for the label is the relationship of the stative meaning to the meaning of the verb as a whole: ‘target state’ participles denote a state which is already an internal component of the verbal denotation; ‘resultant state’ participles denote a state that holds forever by virtue of the event in question having occurred, as in the distinction originally proposed by Parsons (1990) for the perfect. Each of these distinctions comes with its own set of diagnostics. In fact, as we will see, it is not the case that target state participles are the non-event implicating participles and resultant state participles are the event denoting participles. Rather, target state participles can be both event implicating and non event implicating in the relevant sense, while resultant state participles are only event implicating.²⁸

For the present monograph, we want to understand the English system first and foremost, and I will only secondarily make reference to other languages. The organization of data that I will propose for English will look slightly different from what would arise from an inspection of either Greek or German. I think this is inevitable, because, as I will argue, there will be differences between the way the diagnostics play out. This is because of the meanings of

²⁷ I will use the term ‘passive participle’ here informally to refer to the use of the participle in contexts where there is no expression of the external argument. This should not be taken to imply that I think the passive participle is a distinct lexical item from the perfect participle.

²⁸ This is similar to the point made in Alexiadou and Anagnostopoulou 2008. In Greek, it turns out that participles in *-menos* can be both target state and resultant state in Kratzer’s terms, but always have event implications. Participles in *-tos* on the other hand do not have event implications. Those authors also make a distinction between event implicating participles that include Voice and those that do not.

the individual morphemes in question and also because of the properties of the English verb *be*. In addition, the notion of ‘blocking’ is going to play out differently in different languages depending on the other forms available in the system.

3.1.1 Event Implications

Another reason why the discussion in the present monograph will inevitably be somewhat different from the discussions in the literature so far is that I have proposed a rather different semantic ontology which has a profound effect on how we interpret the notion ‘event implications’. Recall that in the present work, I have argued that the lowest phase in the syntactic hierarchy consists of a domain where elements of D_μ , actual linguistic items are concatenated which contribute to the conceptual content of the situation being asserted by the speaker. These meaning contributions are event properties that are abstractions/generalizations over space and time, and so are partial descriptions based purely on immediately sensory and cognitive classification/judgement. These are still in some sense ‘properties of events’ and contain qualitatively different information from properties of objects, or static relations, but the structures built up at this level do not yet entail the *instantiation* of any particular situation. In this context, then, we must rethink what we mean when we say that a form has ‘event implications’ or not. Let us consider first the following diagnostic, as proposed by Embick (2004).

(2) (a) The door was built open.
(b) *The door was built opened.

According to Embick and much subsequent work, the problem with (2-b) is that the state of being ‘opened’ simply cannot be true in the world unless there has been a prior event of ‘opening’, i.e. it is not something that can be one of the door’s properties before anything has happened to it. In the case of English, the fact that *-ed* is often used for both event implicating and non-event implicating forms is potentially confusing, which is why the few cases where there is an underived adjective such as *open* to use as a clear unambiguous point of comparison are useful for separating the readings. When there is no corresponding underived adjective to ‘block’ the use of the participle, it does appear that a participle in *-en/ed* can be used to give the non-event implicating reading.

(3) The door was built closed.

What is important to note however is that the criterion of event implications in the literature refers to the property of ‘entailing the previous instantiation of an event particular’, not that it has conceptually eventive properties.

But the diagnostics for ‘actual event implications’ are actually not so clear. It has been claimed that temporal modifiers like *recently* diagnose the existence of a previous event because it explicitly locates the event in an immediately previous time interval. However, the test needs to be applied carefully because even adjectival statives like *open* can sometimes cooccur with such adverbials, in which case *recently* only asserts that the state held at some time in the recent past.

- (4) (a) The recently opened door.
- (b) The recently open door .

The existence of non-temporal modifiers that refer to other parts of an event description are even murkier in the context of this monograph. *By*-phrases, instrumentals and manner adverbials could be thought of as diagnosing conceptual subparts of an eventive description, but do not in and of themselves require instantiation in space and time. It turns out that some of these are much better than others when it comes to stative participles. Regardless, Embick systematically argues for a distinction between stative participles with event implications (his ‘resultative’ participles) and stative participles without event implications (the misleadingly named pure ‘stative’ participle).

Above, I repeated examples from Embick (2004) showing the participle in attributive position. The assumption is that the attributive position is a ‘pure stative’ position, but in fact the situation is much more complicated since attributive participles seem to show a wider range of modifical possibilities in some languages than the ‘pure stative’ assumption would suggest (see Lundquist 2008 for some discussion).

If we put aside the attributive position since it is not my central concern here, the problem we face is that it is quite difficult to test the Embick-ian ‘resultative’ participle in English because it is quite systematically degraded across the board in predicative position. In other words, once you put the participle in a stative context, say, following the present tense of the verb *be*, and add the event-implication forcing adverb ‘recently’ (keeping the relevant reading in mind), the results are marginal at best. To my ear the examples in (5) are ungrammatical, and 10 years of reading these examples in the literature have not improved them for me. In (5-a) we see the stative context triggered by the present tense of the verb *be* in (5-b) the stative context is the complement position of *turn out* (after Hallman 2009a), and in (5-c) we have the universal

reading of the perfect. (5-d) shows that the universal reading of the perfect is of course felicitous with the non-event implicating version of *closed*.

- (5) (a) ?The door is recently/carelessly closed.
- (b) ?The door turned out to be recently/carelessly closed.
- (c) ?The door has been carelessly closed since Monday night.
- (d) The door has been closed since Monday night.

It has been acknowledged in the literature that these examples are degraded in many cases, but the excuse given is that the degradedness of the resultative participle arises only because of the necessity of coercing activity verbs into having a salient result state in context—a kind of ‘job done’ reading. This indeed seems to be the case for German, as reported by Kratzer. However, it is only plausible for English if we restrict ourselves to data with verbs like *hammer*. The examples chosen above with *closed* demonstrate that a verb with a perfectly respectable result state is still degraded in the stative resultative in English.

Kratzer distinguishes between ‘resultant state’ participles and ‘target state’ participles, but as Embick (2004) points out, the phrasal target state reading that she analyses and gives a denotation for has event implications of necessity, since it requires existentially binding the davidsonian event variable corresponding to the verb. It is only a pure adjectival reading that corresponds to the ‘pure state’ reading in Embick’s terms. The denotation for the adjective *cool* vs. the target state *cooled* from Kratzer, cited in Embick (2004) is given below in (6).

- (6) (a) *cool*: $\lambda x \lambda s[\text{cool}(x)(s)]$
- (b) *cooled*: $\lambda x \lambda s \exists e[\text{cool}(x)(s) \wedge s = f_{\text{target}}(e)]$

The resultant state passives in Kratzer (2000), on the other hand, are the ones where there is no readily available state in the denotation of the verb’s meaning. Instead, the state that the participle denotes is the state that Parsons (1990) calls the ‘resultant state’. The definition from Parsons is given in (7)

(7) **Resultant states**

“For every event e that culminates, there is a corresponding state that holds forever after. This is “the state of e ’s having culminated,” which I call the “Resultant state of e ,” or “ e ’s Rstate.” If Mary eats lunch, then there is a state that holds forever after: The state of Mary’s having eaten lunch”

Kratzer's diagnostic to distinguish resultant states in this sense from the others (the target states and the pure adjectives) is the incompatibility with the adverb *immer noch*-‘still’. This is because the definition of resultant state means that the state persists indefinitely for ever after the event is over, and therefore trivially the adverb ‘still’ cannot meaningfully be applied to it.²⁹ The resultant state passives given by Kratzer (2000) are shown in the examples below .

- (8) Das Theorem ist (*immer noch) bewiesen.
The theorem is (*still) proven.
? ‘The theorem is (*still) proven.’
- (9) Der Briefkasten ist (*immer noch) geleert.
The mail box is (*still) emptied.
? ‘The mailbox is emptied.’
- (10) Die Gäste sind (*immer noch) begrüßt.
The guests are (*still) greeted.
? ‘The guests are greeted.’
- (11) Die Töpfe sind (*immer noch) abgespült.
The pots are (*still) washed up
? ‘The pots are washed up.’

Kratzer (2000) ’s semantics for the resultant state does not produce a property of events, but rather a property of times directly. However, it is important to note that her semantics for both the target state and the resultant state require actualization and have real event implications, since for her events are instantiated particulars. So with respect to event implications, target state passives and resultant state passives are on a par. The only difference is the way in which that state is constructed.

Note that the semantics of the resultant state is claimed to be the same as the one that we would find with the perfect in English (and German), and indeed it is inspired by Parsons (1990) analysis of the English perfect. But the English ‘resultant state’ construction cannot be built so easily along these lines, as shown by the fact that direct translations of the German sentences above into English are ungrammatical/degraded, while they are completely natural using the English perfect of the passive (*The mailbox has been emptied*, *The pots have been washed up*) as Parsons would expect.

²⁹ The test is not perfect, because, as Kratzer explains, the failure of felicity of ‘still’ could also be due to the fact that the target state is not reversible. For failure to combine with ‘still’ to truly diagnose a resultant state, one must exclude the possibility that it fails for trivial real world reasons.

In German also, the perfect passive in (13) and the resultant state passive in (12) are strikingly similar in meaning.

- (12) Das Theorem ist bewiesen
The theorem is proven
'The theorem has been proven.'
- (13) Das Theorem ist bewiesen worden
The theorem is proven gotten
'The theorem has been proven.'

However, Kratzer (2000) claims there is a subtle difference between (13) and (12) in that the resultant state passive is compatible with a reflexive interpretation in principle (as can be seen if a verb such as *waschen*-'wash' is compared in the two constructions) while a perfect passive is not.

In German, there is another surprising feature of the resultant state passive, which is its complete incompatibility with statives and the requirement of coercion with activities.

- (14) *Dieses Haus ist besessen
This house is owned
'This house has been owned.'
- (15) ??Die Katze ist schon gestreichelt
The cat is already petted
'The cat has already been petted.'

In English, the perfect of the passive is perfectly good with stative verbs, so the semantic constraints on the building of the perfect of the passive have to be different for states in English. But we know already that the English and German perfect differ in the existence of the universal perfect reading which arises in the case of states. On the other hand, a 'resultant state' reading is also possible for states in English, and there is nothing logically wrong with building that kind of reading from stative verbs as far as the semantics is concerned and no coercion seems to be required in the perfect. So it is *still* surprising that the 'resultant state' passive is bad in German for states. Further, according to Kratzer (2000), (15) is quite bizarre out of the blue in German even with activities, and needs to have a 'job done' kind of context explicitly imposed on it to be fine. Once again, the perfect of the passive in English is perfectly good and there is no problem with building a resultant state semantics from an activity verb in principle (this in fact was the motivation behind the Parsons

analysis in the first place since the English perfect gets classic resultant state readings with activity verbs, with absolutely no coercion required).³⁰

I conclude that the semantics that Kratzer (2000) gives for the resultant state passive in German, shown below in (16) is actually too weak.

(16) *Kratzer's 2000 Semantics for Resultant State Participles*

Stem+object: $\lambda e[\text{prove}(\text{the theorem})(e)]$

Stativizer: $\lambda P \lambda t \exists e[P(e) \wedge \tau(e) < t]$

Output: $\lambda t \exists e[\text{prove}(\text{the theorem})(e) \wedge \tau(e) < t]$

This denotation is too weak by itself because it generates a well-formed output as long as the run time of the event, given by $\tau(e)$ has some final moment. It is precisely what we need for resultant state readings of the perfect in English, but it should predict that *all* eventualities should be input to it without any problem or coercion, just as long as the activity has terminated, or the state in question has ceased to hold, contrary to fact.

The Target State passive on the other hand, the one that is compatible with *immer noch*-'still', really does have a strong constraint imposed on it. For these to be formed, the verb in question must contain a caused result state in its denotation. Kratzer (2000) diagnoses this by the fact that a 'for-'phrase is felicitous as a measure of the duration of that caused state. It is precisely these verbs that form good target state passives with *immer noch*, Kratzer argues, that consist of an activity portion and a final state. It is this 'final state' that ends up being the denotation of the formed up participle.

Kratzer (2000) herself argues against doing this via syntactic decomposition, because she perceives a mismatch between morphological constituency and the results of the tests diagnosing the existence of a target state. I will depart from her (and in this be more similar to the DM approach) and assume syntactic decomposition of these verbs into ProcP (activity portion) and ResP (caused final state), on the basis of the linguistic tests, regardless of morphological make up.³¹

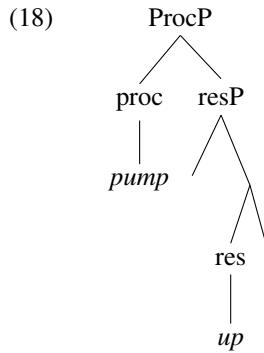
³⁰ I assume that the simple 'perfect' in German is grammatical as well, but since that has taken over the function of the past tense in many dialects, it is plausible that this does not have the same 'resultant state' semantic analysis as the English perfect anyway (see Löbner 2001 for a discussion of the semantics of the German present perfect).

³¹ It is a separate question how the decompositions proposed match up with individual morphemes, particles and suffixes. In the case of German, it seems as if the derivational suffix that derives the verb *leeren*- 'empty' from the corresponding adjective, does not give rise to a target state, while the abstract prefix *auf* does. I assume this is because the causative suffix is actually located in *Init*, while the prefix *is* the licensor of *res* despite its rather abstract semantics in this particular verb.

If we consider the denotation Kratzer assumes for the target state verb *auf-pumpen* ‘pump up’, we see that it contains the representation of a caused final state.

(17) *das Boot aufpump-* - ‘pump up the boat’
 $\lambda s \lambda e [\text{pump}(e) \wedge \text{event}(e) \wedge \text{inflated}(\text{the-boat})(s) \wedge \text{cause}(s)(e)]$

I will represent the existence of a caused internal state as an embedded resP under proc, following the general theory of decompositions proposed in Ramchand (2008). Thus the tree structure licensed by a verb that gives rise to a target state participle, looks like (18).



According to Kratzer’s semantics, the output of the stativizer *-en/ed* is a predicate of states, exactly the one that is inside the verb’s complex event semantics. The external event variable (the ‘process’ variable in my terms), is existentially bound.

(19) Stativizer: $\lambda R \lambda s \exists e R(s)(e)$
Output: $\lambda s \exists e [\text{pump}(e) \wedge \text{event}(e) \wedge \text{inflated}(\text{the-boat})(s) \wedge \text{cause}(s)(e)]$

So the participle morphology in her system does not do very much work except to existentially bind the ‘davidsonian’ event, and also to licence the absence of verbal inflection. My own analysis will be very similar to this idea, but without any actual existential binding of event variables. I implement my version by associating the participial form with a *subportion* of the phrase structure potentially determined by the verb. When it does so, the element of D_μ so formed has a reduced conceptual contribution.

3.1.2 An Implementation in Terms of Reduced Spans

Since the implementation in terms of spanning is somewhat non-standard in the literature, it is worthwhile emphasizing the intuition behind the implemen-

tation first. As already proposed in Ramchand (2008), the lexical entry of a particular verb contains specific information about the subevents it denotes. In fact, in the implementation proposed in this monograph, the LI has a syntactic ‘span’ as one component of its identity, and complex elements of D_μ are built up by merging those syntactic contributions to make derived elements of D_μ .³²

Here for concreteness is a representation of the verb *destroy* in English, which has a syntactic representation $\langle \text{init}, \text{proc}, \text{res} \rangle$ and a semantic contribution as shown in (20).

(20) The verb *destroy* has the denotation:

$$[[\text{destroy}]] = \langle \text{destroy}, \langle \text{init}, \text{proc} \rangle, \lambda e[\text{destroy}(e)] \rangle$$

Further, if we subscribe to the decompositional approach proposed in chapter 1, and necessary here to account for the constraints on target state formation, then we need to decompose the semantic part of representation in (20) into the relevant subevents. Thus, instead of (20), we should have something like (21).

(21) $[[\text{destroy}]] = \langle \text{destroy}, \langle \text{init}, \text{proc}, \text{res} \rangle, \lambda e \lambda e_{\text{init}} \lambda e_{\text{proc}} \lambda e_{\text{res}} [e = e_{\text{init}} \rightarrow [e_{\text{proc}} \rightarrow e_{\text{res}} \wedge \text{destroy}(e_{\text{init}}) \wedge \text{destroy}(e_{\text{proc}}) \wedge \text{destroy}(e_{\text{res}})] \rangle$

In other words, if we have an element in D_μ which is syntactically specified with these three subevental category heads, then the item in question provides conceptual content to each of the subevents in the decomposition. Recall that the central proposal for the event conceptual domain is that these conceptual properties are sensory and cognitive generalizations over experience. The natural hypothesis is that these particular event properties are those that are ascribable based on immediate and direct observation. So, I can look at the lego installation my son has been working on and characterize it as a ‘destroyed city’ based on its visual aspect and my background knowledge of the world. And I can do this even if I have not seen it being destroyed. It also may be the case that my son built it up that way, i.e. he built it ‘destroyed’ as it were. That is the joy of lego and the joy of participle use. It is crucially relevant for my assertion that it is a ‘destroyed city’ and not a ‘flattened city’ or a ‘bombed city’ because those present subtly different aspects. Basically, the verb ‘destroy’ contains information about what the result subevental state should look like, and I exploit that when deploying it to characterize an object in my visual

³² In the normal case also, the finite verb inflects for tense and agreement information, and I will assume that this means that the inflected forms actually include Asp and T in their span information. But we will not be concerned with these syntactic features in this chapter directly.

field. Verbs that have *res* in their lexical specification, have conceptual content that characterizes a state as a having the properties of a final state of Verb-ing.

We do not require a previous destroying to deploy the target participle ‘destroyed’, but it is still eventive in the sense that it makes reference to our conceptual knowledge of destruction events in characterizing that final state. It does not help either to try to express this intuition in terms of some kind of intensional calculation over inertial or prototypical worlds. This fact about how humans deploy verbal lexical items is a basic primitive that is fed by a complex, learned, but immediate judgement. Sensory/cognitive generalization as an abstraction over the space-time reality of experience is the whole basis for the reusability and efficiency of language.

So, the idea is simply that the participle in *en/ed* is the spell out of a *subpart* of the structure listed in the verbal specification. Instead of existential binding of event variables, we simply drop the non-expressed elements of the syn-sem representation. Thus, in the case of *destroy* above, the formation of a target state participle such as *destroyed*, would give rise to a derived element of D_μ which looks as in (22) and which denotes therefore just a simple state.

$$(22) \quad [[\text{ destroyed }]] = < \text{ destroyed}, < \mathbf{res} >, \lambda e [\text{ destroy}_{\mathbf{res}}(e)] >$$

The intuition is that the participial form in *en/ed* is an element of D_μ which is a *subpart* of the related verb root in D_μ , and it is part of its verbal paradigm. The subpart notion corresponds to the Kratzerian intuition that *en/ed* existentially binds off event variables and licences the absence of verbal inflection. In spell-out terms this certainly requires the participial form to spell out a structure that lacks whatever syntactic head it is that hosts the agreement information for tense and subject phi-features in the normal case. But in addition, the claim throughout this chapter will be that the participial form is defined negatively, as the form that is any proper subpart of the verbal root within the first phase. It will therefore be able to spell out chunks of various sizes, as long as the subset property is met.

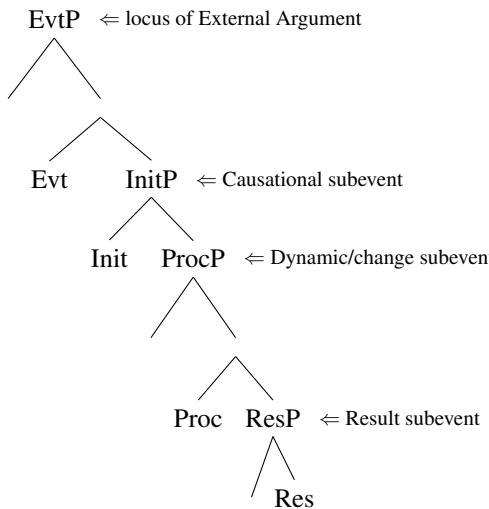
Recall that the decompositional system for the D_μ domain also includes an Evt head within the first phase, as discussed already in chapters 1 and 2. The Evt head is the locus of the introduction of the utterance event and the deployment operator, but otherwise is most similar to the Voice head assumed by many other researchers on argument structure since it will also be the locus of the external argument. I will assume following that work (and against the assumptions made in Ramchand 2008) that the INIT head introduces the causational subevent but not an external argument, while EVT is the head that introduces the external argument when it exists (following original proposals in Pylkkänen 1999).

In recent work on Hiaki, Harley (2013) shows convincingly that whatever morpheme expresses causation in the verbal stem must be hierarchically lower than the head that introduces the external argument. Within her framework, she concludes that Voice and Cause/little v must be distinct, and that Voice embeds little v. In this monograph, I conclude essentially the same thing, with the difference that Evt, the projection that hosts the ‘external argument’ should not be labelled Voice but something more abstract, since the progressive is embedded within it (as argued in chapter 2).

As in Ramchand (2008) however, I will assume that these heads can be specified to be either raising heads or not (i.e. for external or internal merge), and this will have different effects corresponding to the different ‘flavours’ of voice proposed in the literature.

The picture we are left with is thus shown in the tree in Figure 3.1

Figure 3.1
Event Decomposition in the First Phase



The Ramchand (2008) classification of verb types in terms of event structure would then predict entries for the different verb types as follows. In the entries that follow, ‘raising’ heads, i.e. those stipulated to be filled by Internal Merge are notated with a subscripted ‘i’. Differently from Ramchand (2008), the argument introduced in the specifier of *init* is now introduced in the specifier of Evt. Transitive verbs are a heterogenous class from an event structure point of view, depending on whether *res* is present or not.

(23) (a) **Unaccusative**: $\langle \text{Evt}_i, \text{proc}_i, \text{res}_i \rangle$
 (b) **Unergative**: $\langle \text{Evt}_i, \text{init}, \text{proc} \rangle$
 (c) **Transitive**: $\langle \text{Evt}, \text{init}, \text{proc} \rangle$
 (d) **Transitive(result)**: $\langle \text{Evt}, \text{init}, \text{proc}_i, \text{res} \rangle$

Under the spanning view of lexical merge within the lowest zone, the verbal lexical entry is specified for a span of category features that it provides the lexical content for, and can be merged in the first phase to create complex symbolic forms. We first need to be specific about the entry for the tensed verbal form, which I will assume has a different category bundle than the participle in *en/ed*. Specifically, even though the tensed main verb in English stays low by all the usual diagnostics, it does in fact inflect for both tense and agreement, properties of the higher situational domain. I will assume therefore that the verbal entries for tensed roots contain in addition Asp and T features.³³ At any rate, the ‘tensed’ verb in English must be allowed to span all the way up to Asp (via the phase edge Evt). I will use Asp a general category label for all syntactic heads within the situational domain and it will be my name for the domain in which temporal and locational information is added to the conceptual event description. Thus, the entries for the tensed verbs in English actually look as in (24). Each of (a) (b) and (c) below represent a linked cluster of entries with corresponding to the different values of the tense and agreement information, unified under the following categorial specifications.

(24) (a) **Unaccusative (Tensed)** : $\langle \text{Asp}, \text{Evt}_i, \text{proc}_i, \text{res}_i \rangle$
 (b) **Unergative (Tensed)** : $\langle \text{Asp}, \text{Evt}_i, \text{init}, \text{proc} \rangle$
 (c) **Transitive (Tensed)**: $\langle \text{Asp}, \text{Evt}, \text{init}, \text{proc} \rangle$

Now the idea is that the participial form for a particular root will be the form that spells out any verbal span that does not contain the Asp feature. The participial form is any non-tense information carrying contiguous *subset* of those features.

(25) (a) **Unaccusative (en/ed participle)** : $\langle (\text{Evt}_i,) \text{proc}_i,) \text{res}_i \rangle$
 (b) **Unergative (en/ed participle)** : $\langle (\text{Evt}_i,) \text{init},) \text{proc} \rangle$
 (c) **Transitive (en/ed participle)**: $\langle (\text{Evt},) \text{init},) \text{proc} \rangle$

Further, since in this system category features correspond to the existence of particular subevents and arguments in the semantic denotation, we can state an

³³ In point of fact, the tensed verb in English probably only spans all the way up to Asp. I will assume this means that it does not have a T feature in its syntactic span, but a uT feature which enters into agreement with T.

interpretation rule for the case where the a subset of syntactic event structure is deployed, the corresponding conceptual characterization also not present (radically absent from the denotation). This means that when existential closure kicks in at the edge of the first phase, no subevents that have been ‘reduced’ in this way are entailed to exist.

Thus this means that in our implementation of the generation of the stative participle in *en/ed* I assume the following interpreted tree.

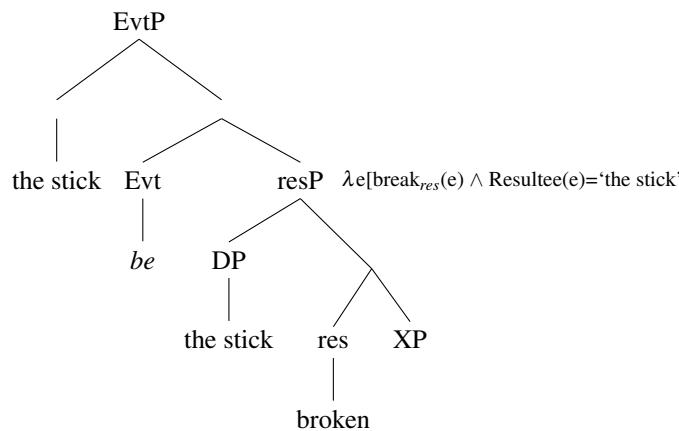


Figure 3.2
Building a Stative Passive Participle

Thus, unlike the Kratzer analysis of the target state however, the implementation as reduced elements of D_μ does *not* have event implications in the sense of requiring an actual situational instantiation of the relevant event. This is in fact the derivation for the productive construction of pure states using *-en/ed*, but without event implications (in the Embick 2004 sense). While these forms do not require a previous actual event instantiation, they do allow some adverbial modification by low adverbs, *if* those adverbs can felicitously modify the result state in question.³⁴

(26) (a) The door was built closed.
 (b) Her hair is still sloppily combed.
 (c) She is still well dressed.

³⁴ The point about certain low adverbs being able to modify target states without event implications is also made by Anagnostopoulou 2003 and Alexiadou and Anagnostopoulou 2008 for Greek.

Even though these forms do not have event implications in the Embick (2004) sense, they *are* crucially derived from verbal meanings and their truth conditions depend on the understanding of the internally complex event that is named by the verb. To this extent, the denotation of the pure stative participle *closed* is different in my analysis from the denotation of a simple adjective such as *open* or *empty* which are not event-dependent meanings. This is because in the new system I have been building, there are two different versions of the idea of ‘event implications’. The first, is a notion that requires an actual event particular to be instantiated in the world. I have been calling event particulars ‘situations’ consistently in this work. The second is a notion that only requires the existence of event concepts as encoded by particular elements of D_μ . Note that the verbal item and its participle in *en/ed* are part of the same paradigm of forms. The analysis of these *en/ed* forms is crucially different from the derivation I assumed for *-ing*, where the suffix *-ing* was its own lexical item, also a member of D_μ . In the analysis of the *en/ed* participle, the ending itself does not have an independent status as a member of D_μ , but is a morphological part of a member of D_μ which is systematically and productively related to the bare verb member of D_μ . This difference correlates with the fact that in the case of the *en/ed* participle the actual form can be idiosyncratic and potentially subject to allomorphy, while the *-ing* forms of verbs are completely regular.

By being linked to the same lexical item, the meaning relationship between the participle ‘closed’ and the verb ‘close’ is in fact not arbitrary, but is systematic and productive. They are connected because they both form part of an event description that has a particular label for the English language user. The recognition of a state as being ‘closed’ is closely tied to the speaker’s understanding of what a ‘closing event’ looks like and what kind of state the object must end up in. In elucidating the truth conditions of the derived adjective ‘closed’, we need to make reference to the idea of closing events—the closing event is primary, and the idea of a ‘closed’ final state is parasitic on it. It simply would not do for the meaning of ‘closed’ to somehow come to mean ‘flat’, or for ‘closed’ to be used of objects that did not have some kind of way in principle of being opened and closed.³⁵ Moreover, stative participles of this kind are formed productively all the time, so their meanings must be productively derivable from the meaning of the caused-state verbs they are based on, and we

³⁵ However, given that the derivation is at the level of D_μ and that a new D_μ is actually built, we predict that it would be possible for them to be reanalysed as not belonging to the same paradigm of forms. If a participle did this, and drifted from its association with the verb it was derived from, it should eventually be cut free and reanalysed by speakers as an independent adjective. I assume this has happened before in the history of English. For example, the adjective *accomplished* in ‘She is an accomplished musician’.

need to be able to do this constructionally since particles and certain final state modifying adverbs can help to license and identify result states in English. We need to be able to express the relationship between the event description and the result state— *even though we do not not require that an actual situation occurred* in order to assert the existence of the result state. So the derivation here is both constructional and productive, but also in some sense lexical since it creates derived members of D_μ .

With regard to intensionality, we are in fact in precisely the same situation that we were in with regard to the progressive, where we needed to be able to assert the existence of an in-progress state without necessarily committing ourselves to the instantiation of the event essence that it was identifying. This kind of ‘pseudo-intensionality’ is the hallmark of the conceptual domain which traffics in abstractions of event properties and where these are manipulated before the existential closure of the event variable being so built up. We need to be able to relate meanings systematically at the level of event concepts and primitive partial descriptions, without committing ourselves to the actual existence of those situations (whether it be in this, or possible worlds.)

The conceptual representation dependent on a verbal root, crucially, is *not* appropriate for underived adjectives such as *open*, which I assume have their own truth conditions independent of verbal meanings, just like any other lexical adjective. However, in these cases in English, it may be that *open* and *empty* are the source of the corresponding verbs *open* and *empty* and *their* semantics is at the core of the verbal meaning, with only underspecified process and initiation components. Their existence furthermore, could be seen to block the use of the *en/ed* participle to express the result state meaning directly.

The other case of participial adjectives that have been cited as not having event implications are the ones like *obstructed* in the following example.

(27) The blood vessel was obstructed.

For Kratzer this is because the participle ending attaches to a verb that has a stative interpretation in the first place, so this is why dynamic event implications go away. This seems correct. But the reasoning does not extend to the pure state participles formed from potentially dynamic verbs like *close*. These do *not* have necessary situational implications.

(28) The box is still closed. It was made that way, and no-one has tried to open it yet.

In my proposal, all of the pure state participles that are formed with *-en/ed* in English and that are compatible with ‘still’ are formed by spelling out the resP

portion of the verbal meaning. They do not have any event implications in the sense of requiring an actualized situational instantiation, but they are consistent with such actualization. The *en/ed* ending does not have a denotation that takes the root meaning as its argument, unlike the *-ing* suffix we saw in the previous chapter. The denotation of the *en/ed*-participle is just a property of states, when it merged and projects only the stative part of the verbal root's denotation. Verbs that have *res* in their meaning, give rise to *en/ed* forms that are *just* a result stative projection. Verbs which do not have a *res* in their lexical meaning cannot form such 'target state' participles.

This makes it a tricky thing to judge the truth of, because the state is being described on its own, but the use of the participle means that it is the kind of state that typically arises from a certain kind of event. I assume that this imposes certain pragmatic constraints on the use of the stative participial construction. It seems that low adverbials often improve the use of these constructions, for some verbs, especially when the result in question is not visible and evident. The discussion of these factors is however beyond the scope of this monograph.

What then of the resultant state readings of Kratzer (2000), or the resultantive passive stative pointed out in Embick (2004)? First of all, as should be clear from the previous discussion, I think that the resultant state passive as described by Kratzer simply does not exist in English. Consider the German sentence in (29) below.

(29) Das Gebäude ist geräumt.
The building is evacuated.
'The building is evacuated/the building has been evacuated.'

Kratzer (2000) describes very carefully the two different readings that exist in German. As a target state passive, (29) means that there is currently no one in the building; in the resultant state meaning, it just means that someone has done their job of evacuating the building, but that tenants might have now moved in again. As far as I can tell, the latter meaning is simply absent from the English sentence 'The building is evacuated'. The meaning that Kratzer is referring to by the resultant state semantics discussed above simply is not a reading that emerges in a stative construction with *be* and an *en/ed*-participle

for any verb in English.³⁶ In English, this meaning is expressed by means of the perfect of the passive.

On the other hand, it is true that activity verbs with *be* do get some sort of possible stative interpretation in sentences such as in (30) below.

(30) (a) The doors are all opened now.
(b) ?The metal is hammered. (c) ?The dogs are chased.

These forms are bad precisely because *only* a Target state derivation is possible for this construction. If a resultant state reading were possible as input to the *be* helping verb, then the sentences in (30) should be *as grammatical as the corresponding perfects*, with the same meaning for the participle. But they are *not* as good as perfects, and they *do* require coercion. The reason is that in order to get a stative interpretation of the participle must express a stative subcomponent of the verb's entry. If the verb does not have a stative subcomponent, as is the case by hypothesis with activity predicates, then the stative participle construction will be bad. The construction can be saved by coercion, but what is coerced is the existence of a contextually available result, where getting the object acted on in this way is an explicit goal of the situation. Specifically, the object must be interpreted as achieving a special new status by virtue of having been 'hammered' or 'chased'. This is a much more special kind of situation than the resultant state semantics of the perfect which is very neutral and requires no special effects on the object. The event just has to be *over* for the perfect to be good (as far as the object is concerned at least). I therefore contend that what is going on with the examples in (30), to the extent that they are good, is that the corresponding verb is being coerced into the structure in (25) with a contextually available resP within the event description. In many cases in order to build this contextual result we rely on contingent facts about the actual instantiated situation (the existence of which is compatible with the derivation in (25) although not required by it). Thus, in many cases, the contextual requirements of the coercion give rise to true actualized situation implications. But the structure is the very same

³⁶ German is known to be different from English when it comes to the meanings and distribution of the verb *sein*, and in the properties of present tense. Present tense in English is simply incompatible with dynamic eventualities, except under special readings such as the habitual, or vivid narrative. My attempts to judge a sentence like 'the dog is chased' are analogous to trying to interpret an event in the present tense, since for me 'be chased' can only be an eventive passive. However, it may well be that German is quite different in this regard, and that there *is* in fact a third category of thing in between the stative passive and the eventive passive. My concern in this monograph is just English, and I intend to make no claims about the potential typology, only about what the system in English actually forms up.

as in the pure state participle derivation. The existence of the extra implications that arise from the process of the contextual coercion, and the fact that true resultant state passives are probably good in some Germanic languages, is what is responsible for the illusion that we have two different constructions here in English. In fact, Alexiadou and Anagnostopoulou (2008) reach a similar conclusion for Greek, at the end of their paper where they find it no longer possible to structurally distinguish between target states and resultant states (in the Kratzer sense). This is because the nature of the state that has to be inferred is actually independent of the structure in their analysis. The structure is the same, but the way in which the resP must be construed is sensitive to the nature of the verb that is being participialized. The same is true of my analysis here. So instead of a resultant state passive stative *and* a target state passive stative, we have *just* target state passive statives, but depending on the verb, coercion is either required or not. The situation is summarized below in (31).

(31) **Participles in *En/Ed* :**

- (a) Stative Participles: Express ResP in a phrase structure.
 - (i) 'Target' State (Verb has res in its lexical specification)
 - (ii) 'Resultant State' (Verb has no res, ResP is coerced and added constructionally)

Neither of these two versions of the target state stative has genuine situational entailments in terms of actual event particulars, but contextual coercion gives rise to situational implications in the case of activity verbs.

Now, the standard DM implementation for the presence or absence of event implications is to include a little v categorizing head in the former case and leave it out in the latter. (In Embick's system, the participial morphology is merged as the category Asp). However, as we have seen, there are at least two different notions conflated in the idea of having or not having 'event implications', and the diagnostics originally laid out in Embick (2004) do not all diagnose the same thing.

(32) • **Event Actuality Implications:** An event of the type named by the verb *must have actually occurred* for the state ascription to be true.

• **Conceptual Event Implications:** An event of the type named by the verb *is the type that has such a state type as its result*.

Pure stative participial passives have conceptual event Implications, and can be modified by low adverbials that can be construed as result state modification, in some cases they even make the result phrase more salient and improve the stative passive interpretations. Pure underived adjectives *cannot* be modified

by eventive adverbials in the same way, and they are not parasitic on event concepts for their truth conditions. Event actuality Implications on the other hand arise in a straightforward way for the English perfect (as we will see in the next chapter), and possibly for some statives in German using *sein*. They also arise from the coercive effects of putting non resP verbs in certain kinds of contexts, but this is a contingent fact related to those particular coercions and does not arise from the structure itself.

3.2 The Eventive Passive

We now move to the eventive passive, which is one of the most common auxiliary constructions in English, and which utilizes the very same participial form that we saw in the stative constructions discussed above. In constructing an adequate analysis of the passive in the framework I am proposing, there are a number of known facts and properties that should be accounted for.

First of all, and most obviously, when a verb is passivized, the agent is no longer expressed as an argument and the object is promoted to the eventual subject position. However, the external argument is semantically present in some way and can control into purpose clauses.

- (33) (a) John ate the apple.
- (b) The apple was eaten (by John).
- (c) The apple was eaten to make a point.

Because of these basic facts, the passive voice has been treated in the literature as involving existential closure over the external argument position. In some cases, this function has been ascribed to the *en/ed* ending itself as actually *being* the existentially bound external argument (Baker et al. (1989)).

Secondly, when the eventive passive applies to a verb, it does not change the stative vs. eventive character of that verb: dynamic eventualities remain dynamic (34), and stative eventualities remain stative (35).

- (34) (a) John ate the apple in 30 seconds flat.
- (b) The apple was eaten by John in 30 seconds flat.
- (c) *John turned out to eat the apple in 30 seconds flat.
- (d) *The apple turned out to be eaten in 30 seconds flat.

- (35) (a) John owns that house.
- (b) The house is owned by John.
- (c) John turned out to own that house.
- (d) The house turned out to be owned by John.

In English, passivization is restricted to transitive verbs, which has led to a characterization of its function in the GB literature as the removal of the ability of a verb to assign accusative case (Chomsky 1981).

- (36) (a) The apple was eaten.
- (b) *The leaves were fallen/*There was fallen some leaves.
- (c) *The train was arrived/*There was arrived a train.
- (d) *The man was danced/*There was danced a man.

However, in other Germanic languages, notably Norwegian, unergatives can also be passivized (Åfarli 1989, Åfarli 1992), giving rise to an ‘impersonal passive’ construction as seen in the following examples. This indicates that the agent suppressing function is probably what is more definitional of the passive construction than accusative case suppression. In fact, accusative case seems to be preserved optionally in the impersonal passive construction in Norwegian, as we see from the third example. All the data below is taken from Åfarli (1992)).

- (37) Det vart sunge
It was sung-PPL
‘There was singing.’
- (38) Det vart gestikulert
It was gesticulated-PPL
‘There was gesticulating’
- (39) Det vart sett ein mann
It was seen-PPL a man
‘*There was seen a man.’

Even in Norwegian though, unaccusatives do not passivize, as shown in (40).

- (40) *Det vart falle eit blad
It was fallen-PPL a leaf.
‘*A leaf was fallen.’

For all the Germanic languages, the passive participle, when used in attributive position, modifies the *internal* argument of the verb. This is in stark opposition to the modifying properties of the *-ing* participle analysed in the last chapter—the *-ing* participle modifies the thing that would have been the subject of the simple verb.

- (41) (a) John ate the apple.
- (b) The half-eaten apple . . .

- (c) John ate.
- (d) *The half-eaten man . . . (meaning ‘John had half finished eating’)
- (e) John danced.
- (f) *The well-danced man . . .

Finally, the data from existential constructions and clefting discussed in the previous chapter, are relevant to put on the table here. By those diagnostics, the passive participle in the eventive passive sits squarely inside the lowest event conceptual domain, before event closure: (i) it is lower than the base generated subject position (42-a); (ii) it can be clefted (42-b); (iii) the eventive participle phrase can not be substituted by British English *do*.

- (42) (a) There were two men arrested at the party.
- (b) I thought Mary should have been scolded, and indeed scolded she was.
- (c) *Mary was arrested, and John was done too.

To summarize, the properties that we must account for are placed together in the list below.

Central Properties of the Passive

- A Existential binding of the external argument
- B Preservation of verbal *aktionsart*.
- C Participle modifies only the internal argument.
- D The Passive VP lies within the lowest Event Domain of the clause
- E Passive does not occur with unaccusatives (Germanic), or with intransitives more generally in English

To this list, we can add another couple of desiderata in the context of this work. The first desideratum is that the analysis of the participle in section 3.1 should be unified with the one that builds the eventive passive as well. The second desideratum is that the *be* auxiliary be unified with the one we found in the progressive construction, namely as a non-contentful element that is inserted in the first phase purely to host inflectional features.

I propose that the analysis of the eventive passive that fulfils these requirements is one that involves the expression of a slightly larger subtree than the one expressed in the stative participle by the *en/ed* form. In this case however, the only ‘leftover’ feature is the one corresponding to the generation of the external argument, namely Evt. I will assume that the existence of the *init* (causing) projection guarantees the existence of some sort of ‘agent’ for the

passive event, but because the verb does not actually project the Evt head that will allow the external merge of a DP fulfilling that role, the external argument gets bound by default existential closure. Instead, an independent Evt head hosting *be* must be merged whose specifier is filled by internal merge from the direct object position. This is shown in Fig. 3.3 below (the denotation given for EvtP abstracts away from the introduction of the utterance situation for simplicity).

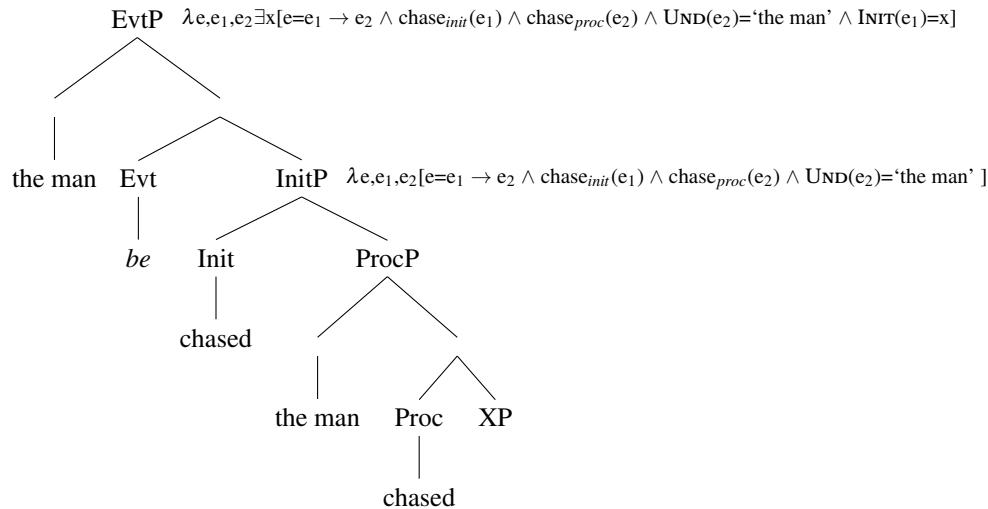


Figure 3.3
Building the Eventive Passive

This immediately gets us properties A-D above. The denotation for the *en/ed* morpheme carried over from the previous section essentially amounts to the existential binding of content corresponding to unmatched features in the participle's denotation. Since the *en/ed*-participle is just an expression of the verbal span up to initP, it does not affect the *aktskift* properties of the verb it is built from. The fact that the only argument that is actually merged at this point is the 'internal' argument means that whatever adjectivalization head is subsequently merged, the argument abstracted over will be the internal one.

How do we then account for the restriction against intransitives and particularly unaccusatives. I will show how this can be achieved next.

3.2.1 Blocking

The restriction in English against passivization of intransitive verbs must be split up into two separate cases. Because Norwegian allows unergatives but

not unaccusatives, the reason for the prohibition must be different in the two situations.

Consider first the case of unergative verbs such as *dance* in English. Nothing in principle should prevent *danced* from spelling out InitP here too as proposed above. However, under the assumption that the lone argument of *dance* is base generated as a specifier of EvtP, this argument will be existentially bound off at the Event domain and effectively removed from the further syntax, and this will leave no argument at all to raise to subject position. English sentences require an overt subject, as is well known (43-a). However, expletive insertion also fails in these cases (43-b).

(43) (a) *Was danced.
(b) *It was danced.//*There was danced.

While I have no deep explanation of this fact, it needs to be English-specific, since as we have seen, it does not hold in Norwegian. I will make the assumption that in addition to the overt subject requirement, English has an ‘EPP requirement’ (in the descriptive sense) at the level of the first phase. For us, this means essentially that there must be an overt DP in the specifier of EvtP. In other words, EvtP requires an overt Topic argument in English. Expletives are not available by hypothesis in the Event Domain, so the requirement amounts to the constraint that there must be at least one event structure argument available in English to construct a grammatical proposition.

The failure of unaccusatives to passivize on the other hand, must have a different sort of explanation. Consider what would happen if the *-en/ed* participle expressed the ProcP determined by an unaccusative verb such as *arrive*.

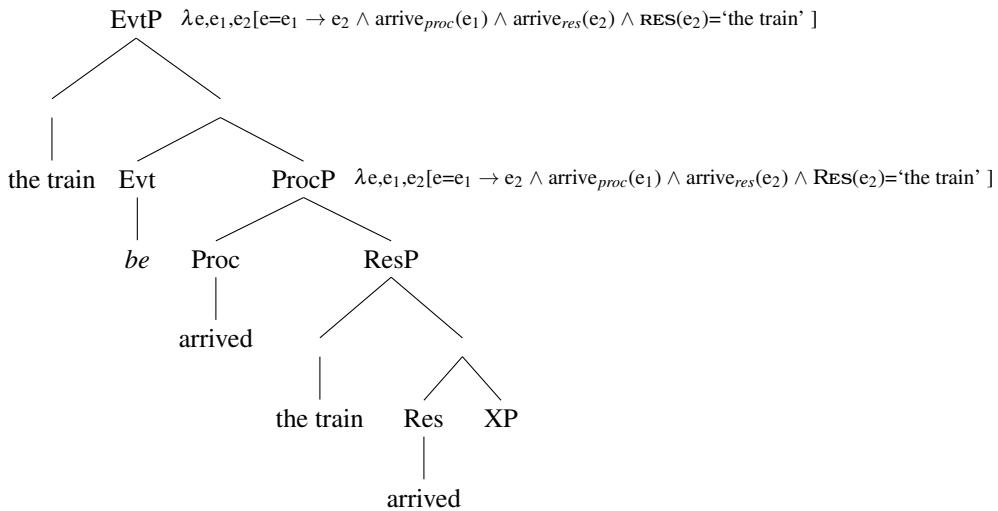


Figure 3.4
Eventive Passive of an Unaccusative

This should in principle be a fine contribution for the participle given what we have said so far, and there would be an argument available to Spec, EvtP. There would simply be no extra external argument which would undergo existential binding. Why then is the passive construction so formed no good (44-a) , although the corresponding stative passive is marginally acceptable (44-b) ?

(44) (a) *The train was arrived.
 (b) ?The train is newly arrived at the station.

One natural thought is that the problem might be a kind of prohibition against vacuous quantification. However, I think that we must reject this possibility. My reasons for thinking so are the fact that the structure above is perfectly legitimate as the input to adjectivalization. There, the effect of participialization is not vacuous presumably because it suspends the continuation of the verb to tense inflection and anchoring, and makes adjectivalization possible.

(45) The recently arrived train....

Indeed we have already seen something similar in the internal distribution of the *-ing* participle which was different in predicative and attributive positions. Having said that, the data are tricky to interpret in English because many verbs that qualify as unaccusative actually undergo the labile causative alternation. There are not many ‘unaccusative’ verbs that do not, and this makes it hard

to be sure that there is not something independently going wrong with eventive passivization for these verbs. In the case of unaccusative *melt* or *break*, the grammaticality of (46-a) is presumably built on the transitive alternant of *melt*³⁷, as is the successful attributive modification in (46-b).

(46) (a) The chocolate was melted over the fire.
 (b) The melted chocolate dripped over the car seat.

But does (46-b) also have an unaccusative source? The eventive passive in (46-a) certainly seems to force the existence of an (existentially quantified) agent, but what about the attributive modification in (46-b)? My intuition here and the judgements of other English speakers I have consulted indicate that the attributive participle can indeed have an unaccusative/inchoative interpretation.

We can look to other closely related languages to complete the argument about where the source of the ungrammaticality should be located. Like English and Norwegian, Swedish does not form eventive passives from unaccusative verbs. However, Lundquist (2008) shows that in Swedish, the passive participle is perfectly good for unaccusatives in attributive position (where no labile alternation occurs to create ambiguity). In (47) we see a passive formed from the transitive version of ‘sink’, and in (48), we see the ungrammatical passive based on the unaccusative. Finally, we see that the unaccusative passive participle is perfectly good in attributive position, and has an eventive interpretation, as diagnosed here by the fact that it is ungrammatical with the Swedish version of ‘still’ (*fortfarande*).

(47) Skeppen blev sänkta
 Ship.DEF was sunk_{tr}-PPL.
 ‘The ship was sunk’

(48) *Skeppen blev sjunkna
 Ship.DEF was sunk_{intr}-PPL
 ‘The ship was sunk.’

(49) Den (*fortfarande) sjunkna ubåten
 The (still) sink-PPL submarine.DEF
 ‘The sunken submarine.’

³⁷ In Ramchand (2008), I argued that in English labile causative/inchoative verbs should be built in the syntax via a null causative head. For the purposes of the system being described here using spanning and its relationship to the participle in *en/ed*, I have to assume that verbs like *melt* actually have an entry with an optional *init* feature. This will mean that the participial spell out of a [proc, res] structure may either have an existentially bound external argument or not. In the latter case, it will be blocked by the simple unaccusative form and will never surface. Only the participle version of the full initP structure will feed eventive passive formation.

This means that there is nothing in principle wrong with a reduction from $\langle \text{Evt, proc, res} \rangle$ to $\langle \text{proc, res} \rangle$. In other words, there is no *requirement* that there be an external argument there to be existentially bound off for well formedness—i.e. the passive is not somehow contributing an existential quantifier that will be give rise to illformedness if its binding effects are vacuous. The existential binding of the external argument we find in eventive passives must rather be some sort of default rule that kicks in when appropriate.

A more interesting possibility for the unavailability of the eventive passive for an unaccusative verb emerges however if we consider what the final passive verb phrase looks like (e.g. in Figure 3.4 above) and compare it to the simple tensed unaccusative tree in Fig. 3.5

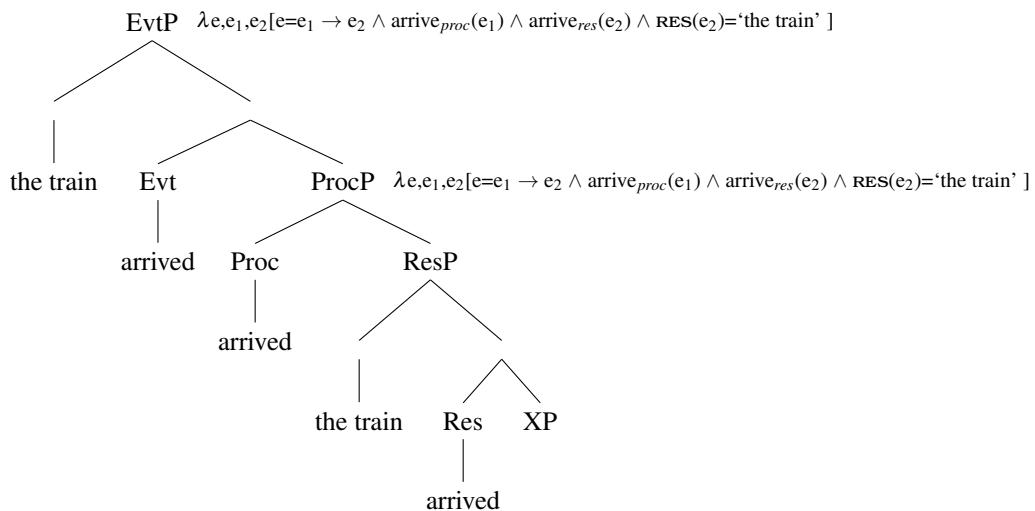


Figure 3.5
Simple Unaccusative

This is identical to the tree proposed for the participial structure both in terms of phrase structure and semantic interpretation. The only difference is that in Fig. 3.5, the structure is spelled out with one lexical item, whereas in Figure 3.4 it is spelled out by the unaccusative participle plus *be*.

Lundquist (2008) also gives a phrasal blocking account for the identical phenomenon in Swedish, and adds to it the evidence from Hindi, taken from Bhatt (2008) which I discuss below.

In Hindi, there is a productive construction by which a simple underived adjective combines with the light verbs *ho*-‘become’ or *kar*-‘do’ to give in-

transitive or transitive dynamic predication respectively. The forms with the adjective *gilla-a*-‘wet’ , are shown below.

(50) kapre giile ho gaye
 clothes.M wet.MPL become go.PERFECTIVE.MPL
 ‘The clothes became wet.’

(51) Atif-ne kapre giile kiye
 Atif-ERG clothes.M wet.MPL do.PERFECTIVE.MPL
 ‘Atif wet the clothes.’

Interestingly, there is another class of adjectives, which are participial forms derived from verbs. These have the same *-aa* ending as the underived adjectives and decline for agreement in the same way as the other adjectives, *but they are systematically ungrammatical in the very same constructional frames!*

(52) *kapre suukhe ho gaye
 clothes.M dry.MPL become go.PERFECTIVE.MPL
 ‘The clothes became dry.’

(53) *Atif-ne kapre suukhe kiye
 Atif-ERG clothes.M dry.MPL do.PERFECTIVE.MPL
 ‘Atif dried the clothes.’

Bhatt (2008) notes the following generalization: forms like (52) above are blocked precisely when there is a simple unaccusative verb, as in (54) and forms like (53) are blocked exactly in the case where verbal forms like (55) exist.

(54) kapre sukh gaye
 clothes.M dry_{intrans} go.PERFECTIVE.MPL
 ‘The clothes dried.’

(55) Atif-ne kapre sukhaa-ye
 Atif-ERG clothes.M dry_{trans}-PERFECTIVE.MPL
 ‘Atif dried the clothes.’

I therefore take the phenomenon of phrasal blocking to be well attested: lexicalization via deverbal morphology plus helping verb is systematically blocked by the existence of underived lexicalization via the simple verb. This is an important enough element in the context of this monograph, to deserve its own label.

Box 3.1
Phrasal Blocking in Auxiliary Constructions

For any two identical phrase structure representations, lexicalization via a deverbal form plus an auxiliary verb is systematically blocked by the possibility of lexicalization by the corresponding simple (underived) verbal form.

We used this fact to account for the absence of stative verbs in *-ing* in chapter 2, and we have used it now for the absence of unaccusatives in passive construction with the *-en/ed* participle here (the latter being attested not just in English, but also in Swedish and Hindi).

3.2.2 Interaction of Passive and Progressive within EvtP

I have now made a proposal for the analysis of the verbal passive, which, like the progressive lies firmly within the conceptual event domain. An immediate question arises concerning how the two interact and combine, and what their relative ordering derives from. Recall that in English the two constructions combine grammatically in only one possible order.

(56) (a) Vidar is being photographed.
 (b) *Vidar was been photographing.

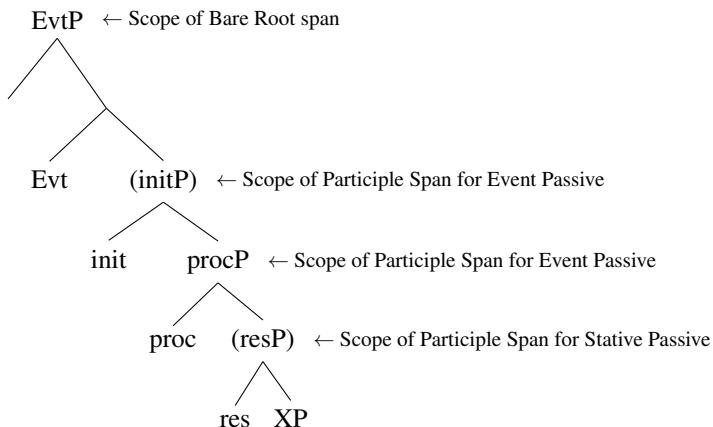
We know that the passive can combine with both stative and dynamic projections, so the fact that the output of the progressive is a state cannot be the impediment to passivization *per se*. In fact, the way that I just posed the question is not correct—we should be asking about the combinatoric properties of our building blocks and trying to derive the fact that the above sentences cannot be generated. So far, I have made a proposal about the height of attachment of *-ing* and the possible expressive structures for the *en/ed*-participle. The *-en/ed*-participle is a building block which corresponds to truncated versions of the verbal root; *-ing* is an independent member of D_μ which attaches to the expression of an EvtP, a complete event description as expressed by a verbal root.

The phrase structure that I have been assuming so far for the lowest domain of verbal concepts contains three force dynamical heads, *init*, *proc* and *res* and an Evt head where the highest argument is merged.³⁸ (I have also assumed

³⁸ Instead of assuming a separate Voice head with a causal flavour, I will assume for concreteness that the Evt head can host derived external arguments and is not necessarily tied to a particular event structure role, although when an argument is base generated there it is interpreted as bearing the CAUSER or AGENT role depending on the encyclopedic semantics of the verbal root phrase *initP*.

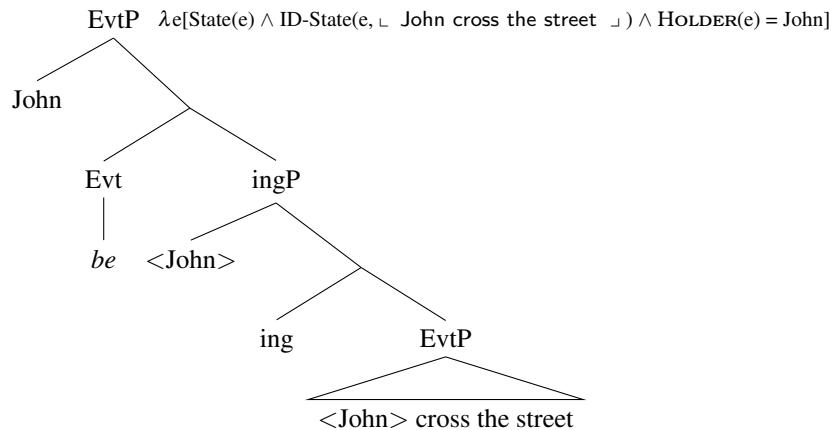
that the Evt head is the locus of introduction of the utterance situation and the deployment operator, but we will be abstracting away from that for the time being, at least until we begin discussing the situational domain and the interpretation of the perfect in the next chapter.) We have also assumed lexical entries containing spans for both inflectable roots and participles. I have assumed that inflected stems contain the full specification of category features of the root plus a feature that allows them to inflect for tense and agreement, the Asp feature. The bare uninflected root on the other hand, contains everything *except* the Asp feature since it does not inflect for tense. Finally the participial forms are subsets of the root form that are contiguous. Comparing the forms, we have:

(57) (a) Inflectable Transitive Verb: $\langle \text{Asp}, \text{Evt}, \text{Init}, \text{Proc}, \text{Res} \rangle$
 (b) Bare Root Form: $\langle \text{Evt}, \text{Init}, \text{Proc}, \text{Res} \rangle$
 (c) Participle in *en/ed*: $\langle ((\text{Evt}), \text{Init},) \text{ Proc},) \text{ Res} \rangle$



In chapter 2, I argued that *-ing* was a head in the lowest event domain as well, and that it creates a derived state from the event description that it attaches to. The span up to EvtP can be spelled out by the bare root form, and this can be followed then by the Merge of *-ing*. Once *-ing* has been Merged and a new event description built, a new Evt head will have to be Merged to provide the edge of the event concept domain and the event concept domain topic for further syntactic action. This will essentially force the insertion of the dummy verb *be* in Evt. In what follows, I will assume the label Evt as a general label for a head in the first phase of the clause, and reserve the label Evt_{edge} for the highest such head in the domain. It is the Evt_{edge} head that has the quotational

semantics motivated in chapter 1, and which closes off the D_μ denotational domain. The successful *-ing* derivation in figure (57) therefore contains EvtP recursion, the hallmark of which is dummy *be*-insertion.



The Merge of the higher EvtP is forced by the requirement of having a verbal form with an Asp feature that will allow tense inflection. I will assume that the lexical item *be* in English has the following specification.

(58) *Be*: < T, Asp, Evt >

Under these assumptions, it is clear that a passive structure then, should be able to feed progressivization because it builds an EvtP (albeit with its own dummy *be*, which should be irrelevant) which as we have seen is what *-ing* attaches to. *Vidar is being photographed* is thus straightforwardly derived. Note that *Vidar is photographing* is not possible in English, for the simple morphological fact that the *-en/ed* participle cannot host suffixation.

Consider again the ungrammatical (59).

(59) **Vidar was been photographing the cat.*

Can this be ruled out in the system we have been constructing so far? To build this sentence, we would first need to attach *-ing* to the EvtP formed by

Vidar photograph the cat. So far so good. Now we would need to insert the *-en/ed* form of *be*. By assumption, this should be the form of *be* that has no *Asp* feature in its lexical entry, and would just be < *Evt* >, something clearly allowed by the rules we have given for participles in *en/ed*. However, because this form does not bear any tense or agreement features, this would then force *Evt* head recursion, once again by assumption, and the Merge of a new *be*, creating (59). The insertion of the participle version of *be* is entirely vacuous here, since the thing it removes when compared to the non participle form, is simply re-added as another token of *be*, this time as a root. I will assume that this sort of derivation with superfluous lexical items is prohibited by some version of lexical economy.

There will however be one case where something like (59) can in fact be built. This will be the case where the truncated form that is the *participle* is actually the expression of everything in the root up to and including *Asp* (but missing the uninterpretable tense and agreement features). In this case, the derivation can only be saved by the addition of an appropriate auxiliary verb in the higher domain, and this will be the story of the perfect.

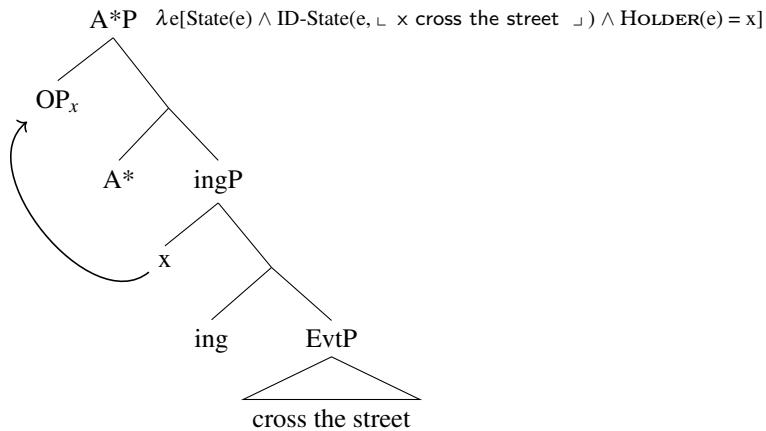
But before we turn to that, we close off this chapter with a brief discussion of *adjectivalization* via *A** of the passive participle.

3.2.3 Adjectivalization

I have given a very simple and abstract denotation for the *-en/ed* participle in its usage as a passive participle in English as an ingredient in both stative and eventive passives. Any span of event concept domain heads lacking an *Asp* specification, and regardless of actual size, can be expressed by the *en/ed* participial form of the corresponding verb. The conceptual content associated with syntactic information truncated in this way is radically absent. In the case of *init* being present while *Evt* is absent, the causing subevent is conceptually present but no actual causer is projected. I have inferred that this is precisely the situation that gives rise to the notional existential binding of a ‘causer’ argument (the presupposition of existence of such a participant), and which licenses *by*-phrases and other agent-oriented modifiers. Modulo truncation, the denotation of the participle in *-en/ed* is claimed to be simply identical to that of the corresponding verb. Depending on the size of the span expressed by the *en/ed*-participle, it is either stative or dynamic, but it has so far lacked an external argument because of the truncation of *Evt*.

So far the selectional properties of the progressive and passive have been derived by combinatorics, and by the existence of blocking with simple verbal forms when in the verbal domain.

But how about the adjectival uses of these participles more generally? How are these built, and how are the restrictions on their usage captured in this system? We saw in the chapter on the progressive that the ingP formed up by the attachment of *-ing* in the Event domain fed adjectivalization. The assumption there was that there was a null adjectivalization head, which I called A*, which induced lambda abstraction over the highest argument in the ingP giving a property of individuals. This derived the fact that in modification, the *-ing*-participle in English always modifies the argument that would have been the grammatical subject if the construction had been input to a verbal progressive. The representation given there is repeated here below.



Now, we know that the *-en/ed*-participle which only ever modifies the argument that would have been an *internal* argument of the related verbal event. Importantly, this is true regardless of whether the argument in question will end up in subject position of the corresponding simple verb or not (60).

(60) (a) John photographed a bear. → The much-photographed bear ...
 (b) John loved a pop band → The much loved pop band ...
 (c) John danced a jig. → *The much danced man ...
 (d) The leaves fell to the ground. → The fallen leaves ...

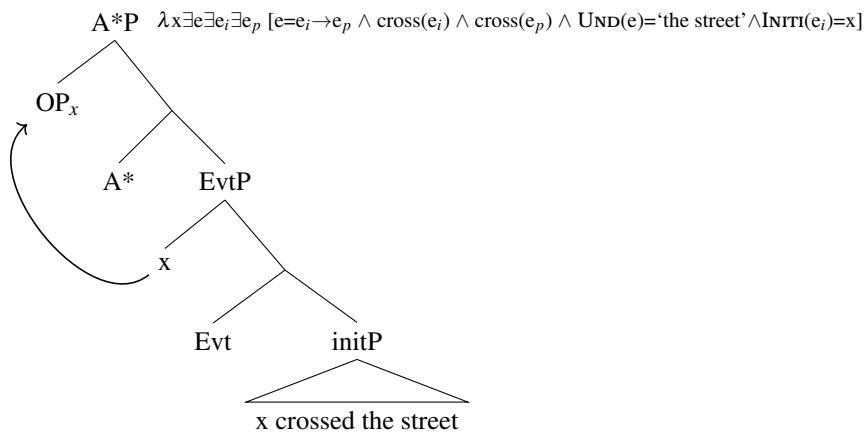
We can now invoke the very same A^* head in the case of the *-en/ed*-participle as well and create an abstraction over the highest argument. This gives ex-

actly the right results for both the stative and eventive passive participles with attachment to *resP* and *initP/procP* respectively.

Note that the prohibition against ‘subject’ modification by bare *en*/*ed*-participles carries over to the attributive use, as seen in (61).

(61) (a) John is photographing a bear. → The man photographing a bear was tall.
 (b) John photographed a bear. * → *The man photographed a bear was tall.

This fact was noted already in Bresnan (1982), stated as the fact that the ‘passive participle’ can be input to the adjetivalization rule, while the perfect participle cannot. In our case, we might ask ourselves why a hypothetical *EvtP* participle could not be input to adjetivalization in this way, on the same model as the participle in *-ing*. (We have blocked its construction in predicative positions on grounds of economy, but as we have seen in the case of the *-ing* participle, such reasoning does not carry over to attributive positions.)



To account for this gap, I will note the stative nature of the *A* category more generally, and stipulate that adjetivalization via *A** in the first phase is restricted to *stative* projections with one designated ‘subject of predication’ position. It can be fed by *-ingP* formation, which is a derived state constructed over the highest position (the other event description and its participant roles are rendered opaque by the derivation of higher secondary Event-state), but

only by *-en/edP* formation if it is based on *resP*. In that case too, we have a state and there is a unique argument in the specifier position.

Recall from the discussion of the stative participle in *-en/ed* in 3.1, I assumed that purely stative participles could be built from *resP* verbs, and also by activity verbs via coercion, if a contextually relevant *resP* could be inferred. These stative predictions had no event entailments in the sense of requiring actual eventive instantiations, but they had event implications in many cases of coercion. We also saw that the addition of adverbials describing the result state were often felicitous, sometimes even facilitating the construction of a result state. I speculate that all instances of derived adjective formation with *-en/ed* are based on the latter attaching to *resP*, either with or without coercion. This predicts that the argument that is abstracted over will always be internal, but also predicts that there may be contextual and pragmatic restrictions on whether the adjectivalizations will be felicitous or not. For example, I think that the addition of the adverbial *well* in (62-b), is a modifier of the coerced result state of hammering. This licences the structure whereby the participle (plus modifier) identify *res* thus feeding stative participle formation.

(62) (a) The window is broken. → The broken window ...
 (b) The metal is well-hammered. → The well-hammered metal ...

The reason that I am able to make the claim that only extremely low attachments of *-en/ed* can feed adjectivalization is assumption about nature of eventive conceptual meaning and their (re)useability. This means that we can allow event ‘dependence’ and felicity of adverbials without assuming full situational entailments.

The claim in this section about *A** is essentially the claim that the null adjectivalizing head has no power to construct a derived state from what it attaches to (adjectives themselves being externally stative). The constituent that feeds adjectivalization must itself already be stative in its denotation, and have a unique argument available for abstraction.³⁹ This means that if the *-en/ed* par-

³⁹ Throughout this and preceding chapter, I have been assuming that we can distinguish in a primitive way between events and states, which then interact in distinct ways in the temporal calculus. We can either encode this as a primitive property of eventuality arguments which can be invoked in meaning postulate statements, or it can be reified as a sortal difference in the ontology itself. I do not know how to distinguish between these two options and I have been simply assuming the former mechanism implicitly, without making the ontology for events more fine grained. It does not seem to me however that the other *aktionsart* categories (i.e. achievement, activity accomplishment) need to be ontologically distinguished in this way since I think their behaviours emerge from the state, event distinction and factors of composition. The state vs. event distinction is a deep and primitive one in this system however. See Altshuler (2016) for discussion and arguments that it should be expressed in the ontology.

ticiples spells out either *procP* or *initP* (or indeed *EvtP*) as we will see in the next section, it may feed further verbalization but never adjective formation. For dynamic verbs, only *resP* *en/ed*-participle adjetivalizations are possible and these will therefore always involve an internal argument. In the case of stative verbs, *-en/ed* must attach low enough so that there is only one unique argument position in order to feed adjetivalization.

Note that the view I am espousing here with A^* is not the same as DM's categorizing little a head. A^* is a head that converts something whose category is already specified as something else into something that behaves externally like an underived adjective. It is not a lexicon-internal device either, since it can attach to phrases and can have quite a lot of internal structure (see also Bruening 2014). I assume that the constraints on adjetivalization are related to the amount of internal structure that is present inside the A^* head. I assume further that stronger constraints on phrasally derived adjectives come either from their lack of scalar properties, or from the coercions independently required to construct a salient result states in different contexts.

Certainly, I do not take external adjetival distribution to be a hallmark of lexicon-internal derivation. In addition, the system I am elaborating here is one in which forms in attributive position and forms in predicative position can in principle have different selectional constraints. This is because although both predicative and attributive uses of the participle have the same participial ingredient, blocking operates differently in the two cases depending on the nature of the embedding structure.

The purpose of this section has been to show how the proposal made here for the participles in the English auxiliary system could be deployed in non-verbal constructions as well. It is however beyond the scope of this monograph to discuss in detail how the pragmatic restrictions are to be stated in each case. That question must be left for further research.

4 The Participle, the Perfect and the Spatiotemporal Domain

In this chapter, I take a close look at the Perfect auxiliary construction in English. It has in common with the previous chapter the deployment of the participle in *-en/ed*. But in other ways we are in new territory because the perfect construction will be seen to implicate the second phase of the clause. We will leave the domain of linguistic lexical items D_μ and be in a zone where spatio-temporal and anchoring properties of eventualities are introduced and modified.

To summarize what we have seen so far, the *-en/ed* participle in English has been argued to express any contiguous subchunk of the verbal root's span, via the non-projection of higher features, and the concomitant suppression of conceptual content. Other than that, the participle simply denotes the property of events that has been built up to that point, as conceptually elaborated by that particular verb. Unlike the *-ing* morpheme which actively builds an Identifying State for the event description it attaches to, *-en/ed* leaves the nature of the event description intact. This has the effect that Target state participles will be stative, while verbal passives will have the same *akitionsart* as their active counterparts. On the other hand, *-ing* never interferes with the argument structure of the projection it attaches to: it attaches to a full EvtP as expressed by the bare root, and always raises the highest argument in that event description to its **HOLDER** position. The participle in *-en/ed* does however potentially affect argument structure because it aborts the expression of higher subevents, changing the options for the choice of externalized argument.

The participle in *-en/ed* is an important one in the construction of derived adjectives, as we have seen, since a stative form with a single privileged subject of predication seems to be systematically able to feed lambda abstraction. So far, I have assumed that this is achieved by the addition of the A^* functional head. This use of *-en/ed* seen so far however, as passive and attributive adjective seems quite distinct from the use of the participle in perfect constructions, an example of which is shown in (1).

(1) Vidar has written the letter ‘V’.

Perfect participles exist with no diminishment of the verb’s argument structure, and they also do not seem to feed adjectivalization, as discussed earlier.

(2) (a) The nicely written letter ‘V’....
 (b) *The boy nicely written the letter ‘V’....
 (c) The boy writing the letter ‘V’...

So even though (2-a) is grammatical, there is no evidence that this is anything other than the passive participle at work. Any attempt to build participial modification over the highest argument of the perfect construction fails. This is in contrast to the progressive participle in *-ing* which does privilege the highest argument of a transitive predication, and which can subsequently feed Adjectivalization, as seen in (2-c). If we now translate directly into Swedish, where the form used in the perfect (the supine) is morphologically distinct from the one found in the passive, it is clear that modification of the internal argument is always achieved via the *passive* participle (3-a) and is ungrammatical with the supine (3-b) Once again, modification of the subject of the verb ‘write’ is impossible even with the supine participle as in (3-c) (B. Lundquist, p.c.).

(3) (a) Den vel skrivna bokstaven ‘V’.....
 (b) *Den vel skrivit bokstaven ‘V’
 (c) *Den vel skrivit ‘V’ pojken.....

Finally, while we saw the formation of the participles in *-en/ed* and also *-ing* systematically interacting with simple lexical forms with respect to blocking, we find no such restrictions on the formation of the perfect in English. *Every* verb in English can form a perfect.

There are thus a number of differences between the perfect and passive participles that might suggest a completely different treatment. The fact that Swedish makes a morphological difference between the two means at least that there is no *necessity* for languages to use the same form for both functions.

(4) **Differences Between Perfect and Passive Participles:**

- Perfect participles never feed adjectivalization
- Perfect participles never reduce argument structure
- Perfect participles are always possible— never ‘blocked’

However, the facts of English strongly suggest that the participle we see in the perfect is paradigmatically the same form that we see in the passive. Allomorphic quirks of the latter carry over without exception to the former. Ideally, we

would like an analysis for the perfect participle that shares a substantial part of its denotation and function with the passive, so that the same vocabulary item can do duty for both.

We have already embarked on a methodology that derives the differences between the *en/ed* forms from the scope of the span they express. Let us see whether this too can be the source of the differences between passive and perfect versions of the *en/ed* participle.

Another thing we should reiterate is that by the tests of height and zone given in chapter 2, the perfect participle and the passive participle are not in the same position. Specifically, the perfect participle comes *before* the expression of the low subject in expletive constructions (5-a) vs. (5-b); the passive participle must come *after* the expression of the low subject (5-c) vs. (5-d).

- (5) (a) There might have arrived many trains at this station.
- (b) *There might have many trains arrived at this station.
- (c) There might be many people arrested at the demonstration.
- (d) *There might be arrested many people at the demonstration.

When it comes to VP fronting, we can once again compare the passive and the perfect directly, even when used separately. The perfect participle does not seem to front very easily (6-a), but the passive one is fine (6-b).

- (6) (a) ??If Mary says that the children will have eaten already, then [eaten], they will have.
- (b) If Mary says that the cakes will be eaten, then [eaten] they will be.

Finally, while British *do*-substitution is crashingly bad for the passive participle (7-a), but marginally ok for the perfect (7-b) .

- (7) (a) *Mary was arrested and John was done as well.
- (b) Mary has written to her local councillor and John has done as well.

To my ear, the VP fronting of the perfect is not completely out, and British *do*-substitution is not completely perfect. The choice of position with respect to the subject in expletive constructions however is a very clear judgement.

It is also true that the *en/ed* participle of a main verb can be embedded under the *en/ed* participle corresponding to *be*, as in (8).

- (8) The boys have been chased.

So when they cooccur, the so-called ‘passive’ participle is clearly lower in the spellout order than the so called ‘perfect’ participle.

What I would like to propose for the perfect participle, therefore is that *-en/ed* in this case spans all the way up to Asp in our proposed functional sequence, just as inflected verbal forms do, lacking only the uninterpretable syntactic features for T and agreement found with the latter.

If the suffix *-en/ed* is allowed to express a span all the way up to Asp, then what we get is the very same property of Events as determined by the verb, including the addition of the external argument, and the position of the participle is higher than the low position of the external argument as diagnosed by the expletive construction test. But now the verb cannot express tense anchoring directly itself, but must combine with a helping verb to receive anchoring information.

I have assumed in the previous two chapters that the dummy verb *be* is an element of D_μ and this can only be inserted in the first phase.⁴⁰ What we will need rather, is the insertion of a functional auxiliary directly in the inflectional domain, the auxiliary *have* which is *not* a member of D_μ .⁴¹

But before I can make an explicit proposal, we need to lay out the background concerning what we know about the semantics of the perfect so that we can have an analysis that captures its core properties, while still keeping the desideratum of unifying the *en/ed*-participial vocabulary item.

4.1 Semantic Background to the English Perfect

The Asp head that I have introduced for transition into the situational domain is named Asp for a reason. The location of this head just outside the first phrase verbal domain makes it equivalent to the functional position that has been the locus for viewpoint aspect in the literature, and which has also been crucial for an explication of the semantics of the perfect, starting with Reichenbach (1947).

That tradition sees the Asp head as introducing a topic time, or reference time, which is a crucial intermediary between the utterance time and the event time in expressing the meanings of the perfect (see Klein (1994) and Demir-

⁴⁰ I am assuming further that lexical entries with a set of category features cannot freely under-associate (contrary to what was assumed in Ramchand 2008), so a tensed form of *be* cannot be inserted directly in T in English.

⁴¹ Myler (2017) attempts to unify *have* in all three of its incarnations: lexical possession verb, light verb, and perfect auxiliary. In my system only the former two can be synchronically related, because of the difference functional and lexical items and the ontological system being proposed here. However, it is natural to think that the structural semantic properties of *have* at least carry over to all uses: stative situation with a single HOLDER argument filled by internal merge. This would have to be the outcome of the way in which grammaticalization works however, rather than a result of these forms being the same vocabulary item.

dache and Uribe-Etxebarria (2008)). In Reichenbach (1947), the perfect tense is characterized by the fact that the reference time is the same as the utterance time, while the event time precedes.

Simple past: ——E,R—————S—————

Present perfect: ——E—————R,S—————

Past perfect: ——E—————R————S————

Figure 4.1

The Reichenbachian View (1947)

The Reichenbachian view has been generalized by Klein (1994) and others to conceive of the Aspect node as being something that can be used to impose viewpoint on the event time by selecting portions of it as the reference/topic time. Some examples of the use of the Asp node to characterize perfective vs. imperfective more generally are given in Demirdache and Uribe-Etxebarria (2008) and Giorgi and Pianesi (1997) to name a few. For example, Giorgi and Pianesi (1997) hypothesise that various tenses are the result of the composition of a relation between E and R (relation 2 in their table) and a relation between S and R (relation 1).

| | | | | |
|-----|-----------------|---------|-----------------|-------------|
| (9) | Relation 1: S_R | future | Relation 2: E_R | perfect |
| | R_S | past | R_E | prospective |
| | (S,R) | present | (E,R) | neutral |

I much of the later work building on Reichenbach, Relation 2 has classically been associated with an aspectual phrase structure node, and specifies the relation between the event variable *e* and the reference time variable *t*. The tense node specifies Relation 1, which anchors the reference time to the speech time (see Klein 1994). The following table in (10) shows the proposal found in Demirdache and Uribe-Etxebarria (2000) for the different aspectual values in Romance and English.

| | |
|------|---|
| (10) | (a) [+Central Coincidence]: (FIGURE within GROUND) |
| | Present Tense: UT-T within AST-T |
| | Progressive Aspect: AST-T within EV-T (b) [- Central, +Centripetal] |

Coincidence]: (FIGURE before/towards GROUND)

Future Tense: UT-T before AST-T

Prospective Aspect: AST-T before EV-T (c) [– *Central*, +*Centrifugal*
Coincidence]: (FIGURE after/from GROUND)

Past Tense: UT-T after AST-T

Perfective Aspect: AST-T after EV-T

The Asp node in the system I am building up in this monograph is in some senses in line with this general pedigree of analysis, but differs from it in others. In the present proposal, recall, in binding the eventuality argument of EvtP, the Asp node introduces a variable of spatiotemporal properties of events anchored in d. To this extent the Asp node must always be present in any phrase structure building a proposition, and it *is* the locus where temporal viewpoint or orientation properties for the event can be expressed for the first time.

$$(11) \quad [[\text{AspP}]] = \lambda f_{<v < v, t >>} \lambda d \exists e [\text{Utterance}(d) \wedge f(d)(e) \wedge \llcorner u \lrcorner(e)]$$

To anticipate, however, I will depart from the Kleinian intuition in arguing that an intermediate reference situation, or topic situation (with a distinct situational variable) is only actually introduced in the context of auxiliary constructions. This will make a clearer distinction between constructions involving modals and perfect auxiliaries (which involve intermediate reference situations) and those like perfective and imperfective aspectual constructions which do not.

For the perfect tense, the consensus indeed seems to be that an intermediate ‘reference’ or ‘topic’ situation seems to be necessary. The central problem of the present perfect is that it seems to say two things at once⁴²: on the one hand the evidence from the tense morphology indicates that a present state is being asserted; on the other, there is an undeniable entailment that a certain event occurred prior to that. The twofold nature of the perfect is what Reichenbach (1947) intended to capture with his notion of Reference time and Event time. But there are questions about how exactly to implement this intuition within the theory. How are the two eventualities related to each other? The different accounts in the literature differ with respect to the position that has been taken with respect to this relationship between the topic situation and the event-related situation in this sense. I summarize the major views below.

$$(12) \quad (a) \text{ *The Resultant State Analysis*:$$

The (present) perfect is a present tense assertion of a situation that car-

⁴² The description of the problem extends to the pluperfect, backshifted to a moment in the past.

ries with it an entailment of a past event (Parsons 1990, Smith 1991, Kamp and Reyle 1993).

(b) *The Indefinite Past Analysis*:

The (present) perfect is an assertion of a past event, with a pragmatic component/presupposition requiring present relevance (Reichenbach 1947, Klein 1992, Inoue 1979).

(c) *The Extended Now Analysis*:

In the (present) perfect, a temporally complex situation is being asserted starting from the past and extending to overlap with the utterance time (McCoard 1978, Pancheva and von Stechow 2004)

Thus, the crux of the matter lies in understanding the relationship between the present state and the event in question. Is that relation purely temporal, or causal and how do the participants in the two situations overlap? To answer these questions we need to take a closer look at the semantic peculiarities that researchers have uncovered to date about the perfect. In what follows, I am indebted to the discussion in Portner (2003).

4.1.1 Aktionsart Sensitivity

Even though what I said about the event preceding the present state is generally true for dynamic eventualities, it is now understood that the English perfect gives rise to a number of different readings with different temporal relations between the ‘state’ and the ‘event’ in question. The interesting fact is that these readings are not just available across the board, but are dependent on the *aktionsart* of the event that has been participialized.

In the *target state perfect*, there is a target state for the verb that still holds at the utterance time. In this case, the present state is clearly *caused* by the event having occurred. Obviously, this reading is possible only for verbs that have a target state to begin with. (Terminology and definition here come from Parsons 1990)

Target State Perfect

- (13) (a) John has thrown the ball on the roof. (and it's still there)
- (b) John has pushed over the chair. (and it's still there.)
- (c) John has broken his glasses. (and they're still broken)

In the *Resultant state perfect*, on the other hand, there is a state that holds for ever afterwards simply by virtue of the fact that a particular event came to pass. This state holds now. The semantics of this state are so weak that

the only thing we can conclude from it is that the event occurred at some time in the past (making it truth conditionally equivalent to an indefinite past tense analysis). Once again, the terminology and definition of this type come from Parsons (1990). The resultant state perfect is sometimes called the *experiential perfect* with which it shares its temporal properties. The experiential perfect gets its name from the fact that the current relevance of the state in question gives rise to an implication that the subject has ‘gained the experience’ of having participated in the past event. It seems that this implication arises very easily in the resultant state perfect. I will discuss the ‘relevance’ facts a little later, and will just classify the event structure relationships in this subsection. Resultant state perfects seem possible across the board, for activities (14-a) , achievements (14-b) , accomplishments (14-c) and even states (14-d). However, for some *aktionsart* types, i.e. activities and events without a reversible final state, it is the only reading possible.

Resultant State Perfect

(14) (a) John has driven a truck (before).
(b) John has reached the top of that mountain (before).
(c) John has broken his glasses (before).
(d) John has lived in Paris (before).

Finally, the other major temporal type of perfect here is the *universal perfect*. In this reading, the state that is asserted to exist in the present is extended backwards to a particular past moment in time and is assumed to hold continuously from that point. This reading is only possible for stative verbs in English, and is always available for them. The universal perfect usually requires a framing interval or *since*-adverbial to trigger the continuous interpretation (otherwise states get a Resultant state interpretation).

Universal Perfect

(15) (a) John has lived in Paris for 3 years (i.e. from three years ago up until now).
(b) John has lived in Paris since 2012.

The ‘resultant state’ analysis found in the literature is designed to deliver the resultant state reading easily, and needs some pragmatic boosting to give the target state reading. It has problems with the universal perfect reading however.

The same is true of the ‘indefinite past’ analyses, which have very similar truth conditions. The most difficult reading to account for under any theory is the universal perfect reading. The ‘extended now’ theories by contrast, are good at accounting for the temporal properties of the universal perfect on the other hand, but need to put in some extra work to explain how the situational complexity of the dynamic readings fits in with out standard theories of tense interpretation.

All of the current analytic options however face the problem of how to get the *akto*sart sensitivity to fall out from a unified definition of the perfect.

4.1.2 Temporal Modification and the Present Perfect Puzzle

The schizophrenia of the perfect can be probed by testing the felicity of temporal adverbials. The present perfect is morphologically marked for ‘present tense’ even though the strongest truth conditional contribution is that of a past occurrence of the event named by the participle. Given an analysis that makes reference to both a reference time and an event time, one might expect to be able to modify either or both with temporal adverbials. The ‘present perfect puzzle’ consists in the fact that in the present perfect (both target state and resultant state readings), the event time resists modification by such adverbials and only the ‘now’ seems to be accessible (16-a) and (b). For the universal perfect, the whole of the time span from a particular past moment to the present one can be provided via a temporal adverbial (16-c), although interestingly the universal perfect reading is not easily available without it and the edges of the time span cannot be separately specified (16-d).

- (16) (a) *John has done his homework yesterday.
- (b) *John has driven a truck yesterday.
- (c) John has lived in Paris for 3 years/since 2012.
- (d) *John has lived in Paris now/3 years ago.

The puzzle becomes even more worrying once one notices that in the pluperfect, both the event time and the reference time do seem to be separately accessible, (17-a) and (17-b) respectively. So, in (17-a), the doing of the homework can have taken place ‘on Thursday’, and (17-b) favours the reading where the topic interval is included in ‘Thursday’.

- (17) (a) John had done his homework on Thursday.
- (b) On Thursday, John had already done his homework.

The event time is also accessible for modification when the perfect is embedded under a modal (18-a), or part of a nonfinite clause (18-b).

(18) (a) John must have done his homework on Thursday.
 (b) Having done his homework on Thursday, John was able to go to the party.

An account for these puzzling facts should ideally emerge from the analysis of the perfect and how the different subparts are linguistically represented. For now, I note only that the fact that the English present tense only combines felicitously with states, either simple or derived, should be a factor in the analysis. We return to this in section 4.2.2.

4.1.3 Lifetime Effects and Current Relevance

Finally, the relationship between the present state and the past event is not as free as the pure resultant state semantics would suggest. The present state has to have some sort of ‘current relevance’ for the perfect to be felicitous. In (19-a), the perfect can be uttered if it is relevant knowledge to know whether John eats raw fish when planning a joint restaurant visit; (19-b) is relevant if we want to know if John is familiar with the city of Paris.

(19) (a) John has eaten sushi.
 (b) John has visited Paris.

In the following dialogues, although the English past tense and the English perfect are often both possible, to my ear (20) is odd in the perfect as an out of the blue description of my day, and I would prefer the simple past. But the dialogue in (21) is fine if I am telling my partner that I have done some exercise and am now ready for a hearty dinner.

(20) A: How was your day?
 B: ??I have swum a whole kilometer today in the pool/I swam a whole kilometer today in the pool.
 (21) A: Are you hungry?
 B: I have swum a whole kilometer today, so yes.

Crucially here, the relevant state seems to be one of newly gained ‘experience’ on the part of the subject. Current relevance is therefore not the whole story here, although it is true in the extremely general Gricean sense of speaking relevantly to your interlocutor. In addition, it seems that some pragmatic work needs to be done to create a post-state for a verb that does not naturally have one. In this case, the experiential effects of the event on the subject seem to come to the fore.

In the case of verbs that have target states built into their meaning, the state in question can always be the target state. But here, we seem to get a kind of evidential constraint on felicity (see also Pancheva 2003). In (22), the perfect is infelicitous if A is interrogating B back at the cabin, even though the tracks are still in the snow, and even though that state is clearly ‘relevant’.

(22) **BACK AT THE CABIN**

- A: How did you find the wounded deer?
- B: The poor animal left bloody tracks in the snow.
- B: ??The poor animal has left bloody tracks in the snow.

On the other hand (23) is perfect if A and B are together in the forest and contemplating the tracks as they speak.

(23) **OUT IN THE WOODS**

- A: How will we find the deer?
- B: No problem. Fortunately, it has left tracks in the snow.

There is also a sensitivity to the nature of the subject or **HOLDER** of the present state, which may or may not be the same as this idea of ‘current relevance’. It has been noticed in the literature that if the subject of the perfect is a historical person, then the perfect is very odd. As noted by Chomsky (1970), (24-a) is rather odd out of the blue, while (24-b) is perfectly fine. (24-c) is an old classic sentence from McCoard (1978), cited also in Portner (2003).

(24) (a) ??Einstein has visited Princeton.
 (b) Princeton has been visited by Einstein.
 (c) ??Gutenberg has discovered the printing press.

However, as Portner (2003) notes, if the discourse topic is understood to be ‘famous people who have visited Princeton’, (24) (a) improves considerably. Portner’s own account is a development of the intuition found in Inoue (1979) that a ‘a proposition expressed by the perfect sentence is ‘relevant’ in that it is in a logical relation to another which is ‘at issue’ in the conversation. In Inoue (1979)’s account, this pragmatic requirement of relevance sits over and above the temporal contribution of the perfect form itself, which for Inoue is a version of the ‘indefinite past’ theory.

According to Moens and Steedman (1988), on the other hand, the perfect is only acceptable if one can identify a sufficiently relevant result, given real world knowledge. Smith (1991) identifies the result state with the *subject* having some relevant property. This involvement of the subject argument goes some direction in helping to account for the lifetime effects noted above—if

a person is no longer alive, they themselves cannot possess any properties any more.

These facts are all relevant to establishing what the asserted state actually is in the case of the perfect and how it relates to the event description contributed by the participle.

To summarize briefly, there appears to be a condition of discourse relevance on the use of the perfect that is often connected to some property of the Subject. In addition, there is an evidential flavour to that stative property in many circumstances.

4.2 The Proposal

My own version of the perfect will follow the syntax and morphology of the perfect directly and build it around the present tense assertion of a situation s' which is necessarily a consequence of the situation denoted by the participle s_0 . Thus, the analysis involves two distinct situations, an embedded one, and another one related to it which will be the essential equivalent of what people have called the reference, or topic situation. The two situations are thus respectively:

- (i) *The Dependent Situation s_0* (the situation existentially closed at Asp)
- (ii) *The Asserted Situational State s'* : the situation introduced by *have* that is in relationship with the dependent situation

The crucial question now is what the ‘Have’ predicate that relates s' to s_0 actually means. The meaning of the perfect needs to include the idea that s' *entails* s_0 , that its existence necessarily entails the existence of s_0 .

For Portner (2003), the relationship between s_0 and s' (in his terms, the pre-jacent event and reference state respectively) is a necessity operator utilizing an epistemic conversational background. Moreover, for him, the perfect presupposes that the reference state (our s' , his q in the formula below) is a partial or complete answer to the discourse topic. Portner (2003)’s version of the presupposition is given below in (25).

(25) A sentence S of the form $\text{PERFECT}(\phi)$ presupposes:

$\exists s [\text{ANS}(s) \wedge \mathbf{P}(p, s)$, where

p is the proposition expressed by S ,

\mathbf{P} is a necessity operator utilizing an epistemic conversational background, and

ANS is true of any proposition which is a partial or complete answer to the discourse topic at the time S is uttered.

In addition, as we have seen, the relationship between p and q , (i.e. our terms s_0 and s' respectively) is dependent famously on *aktsort*. This is expressed in Reichenbachian terms below from Portner (2003) pg 484.

(26) (a) Mary has read Middlemarch.
 Reference time r = speech time (contribution of present tense) Event time $e < r$.
 (b) Mary has been upset (since noon).
 Reference time r = speech time (contribution of present tense) Event time $e \text{Or } e < r$

The crucial innovation in Portner (2003) is the use of a temporal sequencing principle which acts as a default and which allows us to make the semantic contribution of the perfect itself unified instead of disjunctive. Portner's (2003) temporal sequencing principle (TSP) is given below in (27).

(27) (TSP) For any tenseless clause ϕ reference time r , and event e ,
 (i) if ϕ is not stative: $[[\phi]]_{r,e}$ implies that e precedes r ; and
 (ii) if ϕ is stative: $[[\phi]]_{r,e}$ implies that e either precedes or overlaps r .

Portner (2003) points out that TSP, or something like it, seems to be operative not just here in the perfect but also in embedded clauses, and in discourse sequencing. These are contexts where states and events systematically behave differently when related to other eventualities. For example, consider the 'sequence of tense' sentences below where a past tense attitude predicate embeds another past tense. In (28-a) with an embedded state, the 'believed' eventuality can either precede or overlap with the 'believing'; in (28-b) with an embedded dynamic event, the 'believed' eventuality must precede the 'believing'.

(28) (a) John believed that Mary was ill.
 (b) John believed that Mary won the race.

I think that Portner's insight here is important, and I will use a version of it in my own analysis. The insight boils down to the idea that the difference between the perfect of dynamic events and the perfect of states cannot and should not be built into the denotation of the perfect itself, but should be made to fall out from the way that eventualities relate to each other *independently of the perfect*.

In fact, we already have a version of the core difference between dynamic and stative eventualities built into our constraints on temporal instantiation in chapter 2. Ideally, this very principle should be enough in our case as well to get the *aktsort* sensitivity of the perfect.

Another thing to point out at this stage is that Portner (2003) does not in fact attempt a decomposition of the different morphological/lexical components of the perfect, although he does try to separate the semantic from the pragmatic effects. The agenda in this monograph is somewhat different. We need to give a denotation for the participle, as well as a denotation for the auxiliary *have* that will compose to give the required effects.

I have assumed, as I think is natural, that the *-en/ed* participle is the component that directly contributes the embedded situation s_0 , and it is the auxiliary *have* that introduces the secondary stative situation s' . Our syntactic evidence points consistently to the perfect participle being placed higher than the base position of existential arguments and therefore I have proposed that it must result from the *-en/ed*-participle spelling out AspP.

Have must now attach to this constituent to build the perfect, raising the highest argument to its own specifier position. Recall again the proposal for the AspP head given earlier:

$$(29) \quad [[\text{AspP}]] = \lambda f_{<v,v,t>} \lambda d \exists e [\text{Utterance}(d) \wedge f(e)(d) \wedge \perp u \perp(e)]$$

So at the level of AspP we have a property of Relations that link the utterance situation d with an existing event that is being demonstrated/described in d . That event has conceptual/perceptual properties as characterized by u . At this point we must allow *have* to combine with this constituent to build a derived state that will have a particular relationship to e , but which will itself be the actual eventuality that is explicitly anchored to the utterance by means of tense.

As the final ingredient, I need to provide a denotation for the interpretation of *Have* and how it combines with the default AspPs so generated.

The denotation I propose for *Have* is given below.

Box 4.1
Denotation for Perfect *Have*

$$[[\text{have}]] = \lambda Q \lambda x \lambda f' \lambda d \exists s' \exists f [Q(f)(d) \wedge \wedge \text{State}(s') \wedge \text{HOLDER}(s') = x \wedge f = \lambda s \lambda d [s' \text{ gives evidence for the spatio-temporal relation between } s \text{ and } d \text{ in the same world as } s'] \wedge f'(s')(d)]$$

The key to the semantics is the definition of the notion of inference licensing state, or evidential state (what I will call the **EVID-STATE**).

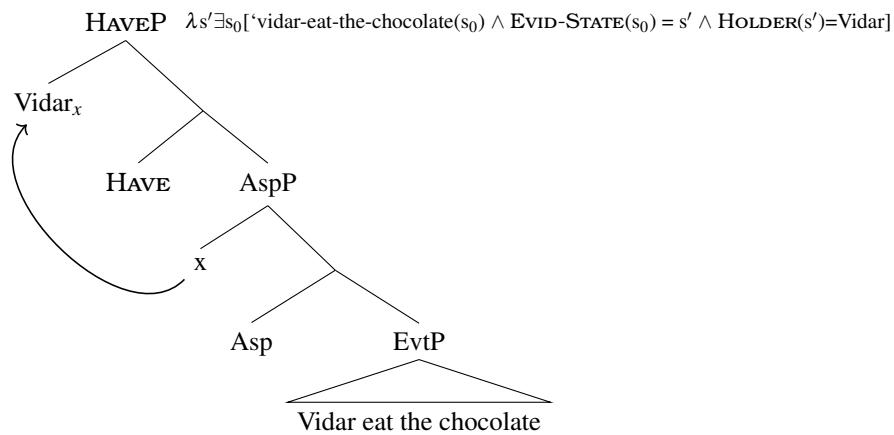
$$(30) \quad \text{For all } s', s, s' \text{ is an EVID-STATE for } s \text{ iff } s' \text{ is a state which gives evidence for } s \text{ in the same world as } s'.$$

By asserting that the f relation between s_0 (our earlier e) and d is facilitated by the existence of s' , I am saying essentially that s' is an evidential state which allows us to infer the existence and spatio-temporal location of s_0 , the event built up by the AspP that *have* combines with. Putting together *have* with the denotation of the AspP (and renaming the eventuality variable s_0 to make the connection with the previous discussions clear) we get the following.

$$(31) \quad [[\text{haveP}]] = \lambda x \lambda f' \lambda d \exists s' \exists f \exists s_0 [\text{Utterance}(d) \wedge f(s_0)(d) \sqsubset u \sqsubset (s_0) \wedge \text{HOLDER}(s') = x \wedge \text{EVID-STATE}(s_0) = s' \wedge f(s')(d)]$$

The **EVID-STATE** s' will have a **HOLDER** in the domain of real instantiated individuals, and its position will be filled by internal merge from the AspP, raising the highest argument there to that role.

In simple terms (abstracting away from the quantificational event semantics formulas) we get the tree below, with the simplified denotation given. At this point, it is the **EVID-STATE** situational variable s' that will be input to modification and tense modulation (anchoring to the utterance) and the embedded situation s_0 will accrue anchoring entailments from those relations indirectly because of its relationship with s' .



The denotation above says that *have* combines with a situational description and creates a *derived stative* situational description, such that the derived stative situation is an **EVIDENTIAL STATE** for that situational description. In the progressive chapter, *-ing* built an **ID-STATE**, which was a relationship be-

tween Event properties, where the identifying state property did not entail the existence of the whole event. Here we are in the situational domain after the existential closure of the event variable. Here, if one situation is inferrable from another another then the existence of the one entails the actual existence of the other.

It is here that the temporal relationship between s' and s_0 becomes crucial. As with Portner 2003, this is the independent factor that is sensitive to *akto*-*sart*. Unlike Portner (2003) however, I will not invoke an explicit Temporal Sequencing Principle. Rather, I only need the principles of temporal instantiation already proposed in chapter 2, inspired by Taylor 1977.⁴³

Here's how the *akto*-*sart* conditions work. Inferring the existence of a situational particular s_0 from s' requires that whole situation to exist *at or prior to* the onset of s' . This is because, according to Werner (2006) and others, only the present and the past are 'determined' in this sense. As we have already assumed, a situation instantiating a dynamic eventuality has a temporal parameter which must be an interval larger than a moment, while a situation instantiating a state only requires that the state have the temporal parameter of a moment. A stative situation can therefore overlap with the stative s' and still be consistent with s' giving evidence for s_0 (because all that is required is a moment). So precisely in the case of dependent states, the s' can in fact perfectly overlap with dependent stative situation (and potentially continue on from there), as in the case of the universal perfect. In the case of dependent dynamic situations, the evidential situation can at best overlap with its final moment or result state, and so the dependent dynamic situation must end up *preceding* it.

Thus, for the assertion of s' an EVID-STATE based on the dependent situation s_0 , we have the following corollaries.

Box 4.2
Akto **Sensitivity for Evidential States**

If f (relating s_0 and d) is inferrable from s' , then we call s' an EVIDENTIAL STATE for s_0 , and then s_0 must be *determined* by the onset of s' .

⁴³ The independent temporal phenomena Portner notes are also relevant in some sense since they indicate the pervasive effects of these general properties in a number of different empirical domains. I assume that the same general mechanisms are at work and give rise to sequence of tense akto-sensitivity and discourse sequencing akto-sensitivity as well. A detailed discussion of these particular contexts and the additional factors those constructions contribute to the final constraints is beyond the scope of this work.

If s_0 is a state, then $s_0 \odot s'$, OR $s_0 < s'$
 If s_0 is dynamic, then $s_0 < s'$

The denotation of *have* as introducing an EVID-STATE based on the dependent situation contributed by the *en/ed* participle plus default Asp, has the immediate consequence of making the perfect entail the actual existence of an event as described by the main verb. It also automatically makes the specifics of the interpretation depend on the state vs. dynamic event distinction for the dependent situation.

In the case of the Swedish ‘supine’ which is the spellout of the participle in the perfect as distinct from the passive, I assume that the supine is the undergrown version of the corresponding tensed formative just like the English participle is, but that it also possesses a distinct LI, which is the truncated version of the Asp-less (non tensed) form of the verb. In a spanning model, in contexts of competition, the item with the least extra categorial material is chosen. This means that it is only in cases of the perfect, which spells out all the way to Asp, that the supine will be selected.

Now I turn to a discussion of how the proposal can account for the other important properties of the perfect, namely the pragmatic effects and the distribution of temporal modifiers.

4.2.1 Accounting for the Pragmatic/Lifetime Effects

Recall that the proposal is that the perfect asserts the existence of a stative situation/property that is ‘held’ by the derived subject of predication (here, the ‘highest’ argument of the dependent situation). There are two features of this derived situation that give rise to very particular pragmatic effects and felicity conditions. One is that the subject of the property must be an instantiated individual at the time the situation is asserted to exist. This is essentially the same intuition found in Smith (1991), updated for the situational analysis. Under the view being explored in this monograph, Events and relations to participants do not require instantiation of either Event description or participant nominal description, but once a relationship is established at the level of situations (Events with temporal and worldly instantiation), then actuality entailments follow the application of existential closure both for nominal and verbal extended projections. I assume that the Smith intuition, and its analogue here, is what is responsible for the infelicity of the examples in (32).

(32) (a) ??Gutenberg has discovered the printing press.
 (b) ??Shakespeare has written Hamlet.

I assume however, if the names of historical characters are construed as abstract subject matter, or labels in a more abstract list then the perfect can become felicitous again, as in (33).

(33) Shakespeare has had a great influence on my life.

Next, we want to account for the intuition whereby the perfect seems to be felicitous only when it provides the ‘answer’ to a salient at-issue question raised by the discourse, as in the analysis of Portner (2003), Inoue (1979) and many others. The problem with the notion of discourse relevance here is that it is hard to falsify, and to the extent that it holds, it is hard to distinguish from general Gricean maxims of cooperation that are independent of the perfect *per se*. So for example, if we consider the discourses in (34) and (35) below, an overt at-issue question concerning either ‘John’ or ‘the ball’ can be answered by a transitive main clause either in the perfect or the simple past. My intuition about (34) is that all the B responses except B''' where the ball is in a subordinate clause can be used in a felicitous answer to A.

(34) A: Where is the ball?
 B: John has thrown it on the roof.
 B': John threw it on the roof.
 B'': It's on the roof.
 B'''?: John hurt his arm throwing the ball on the roof.

Similarly, when A asks a question about ‘John’, my intuition is that both the B and B’ versions are mildly deviant to the same extent. The best answer is of course B''', and B'' is completely strange because it doesn’t mention John at all.

(35) A: What’s up with John?
 B: ?He has just thrown the ball on the roof.
 B': ?He just threw the ball on the roof.
 B'': ??The ball is on the roof.
 B'''?: He hurt his arm throwing the ball on the roof.

The point is that the perfect versions in the above sentences and the simple past versions (B and B') respectively are good or bad to the same extent. This is worrying if the pragmatic presupposition here is supposed to distinguish the perfect from the simple past. In terms of explicit discourse felicity for an explicit question, pronominalization and main clause vs. adjunct clause, seem to be the only factors that are strongly implicated. It seems to me therefore that the notion of discourse relevance does not have much bite here, over and

above general Gricean considerations. Add this to the fact that there are clearly situations where the state in question *does* have discourse relevance but the example is ungrammatical (36-a), and situations where the perfect can be uttered out of the blue (36-b), and it becomes very unclear whether we need to build this into the meaning of the perfect at all.⁴⁴

(36) (a) Gutenberg has discovered the printing press.
 (b) The Orioles have won!

The thing we need to capture is the *difference* between the use of the perfect and the corresponding use of the simple past. In the case of accomplishments, the *s'* that is asserted seems often to be the same as the result state of the corresponding verb. In such cases, the difference between the perfect and the simple past is that in the latter, the result state might no longer hold (37-b). But in the perfect (37-a) the result state is precisely the state that is asserted to exist at the speech time. The result state is in fact a prototype example of an **EVID-STATE** since its existence entails the existence of the dynamic event portion that leads to it.

(37) (a) John has thrown the ball on the roof.
 (b) John threw the ball on the roof.

However, the result state of an eventuality is not the only kind of **EVID-STATE** that could be asserted. In cases where the event description has no result state described within it, as in activities, the **EVID-STATE** needs to be more contextually constructed/inferred. In (38-a), we infer that there is some property acquired by John as a result of his having driven a truck. This can be as non-specific as simply the experience of ‘having driven a truck’, or as specific as ‘knowing something about Jane Austen’s style’, depending on the context. Importantly, because **EVID-STATE** is constructed at the level of situations, it need not be something that is built in to the Event description—it just needs to be something that can be inferred by the listener as existing (in context) by virtue of the fact that the dependent situation exists.

(38) (a) John has driven a truck.
 (b) Mary has read Middlemarch.

⁴⁴ General Gricean considerations of relevance will always apply, but then we need to know exactly what situation/property is being asserted by the perfect to determine whether it is relevant in context. My claim in this section is that it is the **EVID-STATE** that is being asserted and that it holds of a particular participant.

So there *is* an important contextual component to the meaning of the perfect here. The interlocutor must infer the relevant EVID-STATE on the basis of real world knowledge, common ground, and the particular issues under discussion. This is what people call the Experiential perfect, which is a cover term for the kind of situation-based EVID-STATE that needs to be inferred from context. I assume that the Event internal result state, when it exists always counts as giving rise to a potential EVID-STATE, but is only possible for resP verbs. On the other hand the experiential perfect is constructed via situations directly, and is available for all types of verbs (including states), as long as the dependent situation in question is now over and has had *contextual consequences*.

(39) (a) John has driven a truck.
 (b) Mary has eaten sushi.
 (c) The deer has left tracks in the snow.

Finally, we can get a handle on the evidential flavour of the perfect by means of this denotation because it asserts the existence of a criterial or *identifying* situation. Recall the reindeer track examples repeated here in (40).

(40) A: How did you find the wounded deer?
 (a) B: The poor animal left bloody tracks in the snow.
 (b) B': ??The poor animal has left bloody tracks in the snow.

Notice that the B' utterance is clearly *relevant* to answering the question in some sense, but the present tense ascription of a stative property to the reindeer, that of having left tracks in the snow, does not seem felicitous. This is because the perfect in (41-b) reports a criterial state as *evidence* of (41-a). If the event is already completed, as A's question shows, then we already know the truth of (41-a) directly. (41-b) is only felicitous when it is precisely that state that is present and apparent to the interlocutors and not the entailed event itself (by general *Gricean* principles of relevance). Thus, if the deer has not been found yet and all we see is the snow before us, then the following discourse is perfectly felicitous.

(41) A: How will we find the deer?
 B: No problem. Fortunately, it has left tracks in the snow.

The fact that a downstream EVID-STATE is what is being directly asserted, rather than the event itself, is what is directly responsible for the evidential flavour of the perfect that is very salient in certain contexts (see also Pancheva 2003).

I conclude therefore that the proposal of EVID-STATE construction fares no worse on the general pragmatic restrictions on the perfect than many of the other proposals in the literature, and in many cases holds the promise of doing better. A detailed examination of the these constraints however is beyond the scope of the present monograph.

4.2.2 Temporal Modification and the Perfect

Finally, we need to give an account of the curious temporal properties of the perfect if this kind of analysis is going to be convincing. I start with a brief demonstration of what I have asserted so far, which is that the perfect actually involves the assertion of a *stative* eventuality, which is the one marked with tense inflection. The evidence that the perfect has the external distribution of a state comes from the standard tests for stativity in English: (i) it is good in the simple present tense without a habitual interpretation (42-a); (ii) it is good in the embedded complement of *turn out* (42-b); (iii) when following a simple past tense utterance, the past of the perfect does not advance the discourse time in a simple narrative context (42-c); it allows epistemic readings under *must* (42-d).⁴⁵

- (42) (a) John has written a novel.
- (b) John turned out to have written a novel.
- (c) Mary entered suddenly. John had fallen off his chair.
- (d) John must have written a novel.

The present proposal involves the compositional building up of two situations: a first property of (properties of) situational descriptions based on the event built by the *-en/ed* participle, s_0 ; a second situation s' which is an EVID-STATE for s_0 . Only the verb introducing the s' situation is directly tense marked in the perfect. We would expect temporal modifiers to be able to modify this higher situation directly, and in (43-a) and (43-b) it seems to do so, for the present perfect and past perfect respectively.

- (43) (a) John has done his homework now.
- (b) John had done his homework already last Tuesday.

The present perfect puzzle consists in the observation that temporal modification of the embedded situation is possible in the case of the pluperfect, and non-finite forms, but not in the case of the present perfect (44).

⁴⁵ This latter diagnostic will be discussed in more detail in chapter 4.

(44) (a) *John has done his homework yesterday. *present perfect*
 (b) John had done his homework the day before. *pluperfect*
 (c) Having done his homework the day before, John was free to go to
 the party. *non-finite*

To have a consistent description, we could either have a system whereby only the outer situational variable could be temporally modified, OR one in which both situations are accessible for temporal modification. It seems difficult to devise a system where in the present perfect only the outer variable can be modified while in the past perfect and modal perfect the embedded situations are temporally modifiable. Solutions to the present perfect puzzle have included analyses in terms of definiteness of reference time feeding which eventuality descriptions get targeted by temporal modifiers (see Alexiadou et al. 2003 and the introduction therein for discussion). Here, I note that the Extended Now theories (after McCoard 1978) seem to capture an important empirical intuition here.

It is important to note that the description of the present perfect in allowing only present ‘now’ modification, is not strictly accurate. The present perfect in fact also seems happy with adverbials that frame the whole sentence, including both the existence of s' and s_0 (45). Compare the readings of (45-a) with a perfect and (45-b) with a simple state.

(45) (a) John has written a novel this year.
 (b) ?John is tired this year.

A simple present tense utterance of a state is not good with a frame adverbial of this kind, unless the state in question holds *at every moment* during that interval. The reading in (45) on the other hand is that the writing of the novel took place during the period of this year, and that at *some* point during the year John achieved the state of ‘having written a novel’. Data like this are also the starting point from the new proposal in Kamp et al. (2016), where they seek to augment the result state analysis of Kamp and Reyle (1993) to account for these temporal properties that seem at first blush to be at odds with the result state intuition.

Since the compositional reasons for retaining the result state representation for the perfect have not gone away, the solution is to rethink the relationship between temporal anchoring and temporal adverbials. I quote here from Kamp et al. (2016), who say “linking to the time provided by tense and linking to the denotation of a temporal adverb are two distinct operations” (pg 16). The whole ms. is a detailed study showing that a systematic system of adverbial interpretation can indeed be given and that it is perfectly compatible with a

result-state analysis of the perfect, once the assumption given in (iv) below is dropped.

“(iv) The semantics of a projection of the verb takes the form of a single eventuality (either a single event or a single state). It is this eventuality that gets linked both to the time t introduced by tense and to the time denoted by the temporal adverb in case there is one.”

As with Kamp et al. (2016), the analysis of the perfect I have offered contains more than one eventuality. Like them, the s' eventuality is a generalized kind of result state—a state that allows the existence of the embedded eventuality to be inferred and which can indirectly locate it in space and time. Like them also, I propose to account for the curious modification properties by adjusting our assumptions about what is a possible target for temporal modification precisely in the case where a complex eventuality structure is present. Below, I present a brief and informal idea of how that would go for a few core cases for English, but a full treatment of all such temporal modificationary possibilities is beyond the scope of this monograph. I refer the reader to Kamp et al. (2016) for a more complete account.

Ideally, we would like to have conditions on temporal modification of the perfect that are uniform for both the present perfect as well as the past and non-finite instantiations. I propose therefore that the conditions on temporal modification are to be stated quite generally as in (46).

(46) **Temporal Modification of the Perfect**

In the English Perfect, temporal adverbials can modify *either* the run-time of the reference situation s' ((outer)situation-modification) *or* the whole interval containing both the run times of s_0 and s' (frame-modification).

We have seen how this works for the present perfect, but how does it play out in the cases of the pluperfect and nonfinite perfect where it has been claimed that s_0 can be modified directly? The answer comes from the difference between present tense and the other two cases. In the present there is a contextually precise and definite right edge to the interval that is modified in the framing modification alternative (i.e. the utterance time). In the case of the past tense, the reference situation s' does not have to be a point, but can have extended reference starting right after the end of the dependent situation. This is because s' in the case of the past is not identified with the utterance moment, but is a stative situation that occurred *before* the utterance moment. There is thus no requirement that a stative situation s' *must* have a point as its temporal run time, only that it *can*. Consider then the sentences in (47) below.

(47) (a) John had done his homework (already)⁴⁶ at 8 that morning. (*time of s'*)
 (b) John had done his homework the day before. (*time of s₀* ?)

It is usually assumed that (47-a) expresses that the *s'* of ‘having done his homework’ held at 8 that morning, and that (47-b) says that the *s₀* of the ‘doing homework’ event happened the previous day. However, if the *s'* of ‘having done his homework’ can potentially extend right from the moment of the culmination of *s₀* as the denotation of EVID-STATE allows, then the adverbial in (47-b) could in fact be modifying *s'*.

If I am correct that the adverbials that have been treated as modifying the embedded event are actually modifying the reference situation *s'*, then we predict that we should not be able to modify two separate time moments. Under the hypothesis that the pluperfect allows modification of both *s'* and *s₀*, we would predict (48-b) to be grammatical, under the hypothesis that only *s'* can be directly modified but that the choice and location of *s'* is more flexible in the case of the pluperfect, we predict (48-b) to be odd. To my ear, (48-b) is infelicitous.

(48) (a) When his Mum arrived, John had (already) done his homework.
 (b) ?When his Mum arrived, John had done his homework at 8 that morning.

The same effects can be duplicated with the non-finite forms.

(49) (a) Having (already) done his homework when his Mum arrived, John was ready to leave.
 (b) ? Having done his homework at 8 that morning when his Mum arrived later, John was ready to leave.

The compositionally built up complex situation that we call the perfect, therefore has quite regular modification properties, once we take the special property of the present tense into account (i.e. confining *s'* to a single moment). Note that this is the opposite of the claims of the Extended Now family of analyses that say that the present tense can be used to encompass a longer time interval starting from a contextually relevant moment and stretching up to include the Now. However, the intuition of the Extended Now analyses is recaptured in this framework with the stipulation that the *framing* interval for

⁴⁶ In this and the following examples, I use the adverb *already* to force the intended reading. The reading is available without the adverb though.

s' and s_0 is what is available for modification. In the case of the pluperfect and nonfinite perfect, this interval can be described by *by*-phrases, and in the case of the universal perfect, by *since*-phrases (50). *Since*-phrases modify the left edge of the frame and *by*-phrases modify the right edge of a frame.

(50) (a) John had done his homework by midnight.
(b) Having done his homework by midnight, he got at least 7 hours of sleep before having to go to school.
(c) John has lived in Paris since 2005.

When it comes to other adverbials, the current system predicts the constraint on temporal modification above to apply to properties of situations, and does not say anything about other kinds of low adverbials. Specifically, I would assume that event modifiers would be able to unproblematically modify the e that gives rise to s_0 . Since there is no other conceptual event description in a perfect sentence, we do not predict any ambiguity when it comes to instrumental and manner adverbials. Moreover, unlike with temporal modifiers, we should be able to express manner modification of the event at the same time as temporal modification of s' , as (51) shows.

(51) When his Mum came home, John had already done his homework carefully/with an ink pen.

So it is the derived situation s' , the highest one, the one that is actually anchored by tense inflection that can be directly temporally modified, and this is true for all versions of the perfect. It is only apparent/obvious in the present perfect, but it is true quite generally.

I conclude that a treatment in terms of derived EVID-STATES does at least as good a job as the current theories on the market, and in addition builds in the aktionsart sensitivity and evidential properties of the construction. Temporal modification seems to be constrained to the outer situational variable and possibly also total framing adverbials, and it seems like one could tell a story whereby lower temporal modification of the embedded situation alone is systematically disallowed. However, nothing in the overall proposal crucially depends on this, the system as proposed offers the same potentials for modification as others in the literature. I therefore leave a full treatment of temporal modification in the English perfect for further study.

4.3 Summary

In these two chapters, I have set out to give a unified analysis of the *-en/ed* participle in English as it appears in both passive and perfect constructions. The

factors that contribute to making these constructions come out differently are (i) the height to which the participial span extends and (ii) the other functional material that gets merged subsequently. For the passive, the stative and the eventive passive spelled out resP and initP respectively, denoting within the domain of D_μ , while the perfect was built from a participle that was formed from spelling out AspP and included the external argument. More importantly, the perfect participle then required the merge of a verbal head in the situational domain to complete the construction of a proposition. This was *have* and the construction of the derived evidential stative situation gave the temporal and other semantic properties of the perfect that have been described in the literature. Specifically, the fact that this relationship is now built over situations, which are particulars instantiated in time and world, gives rise to the special entailments for the participial description that we did not find when we were still manipulating objects of the D_μ semantic type.

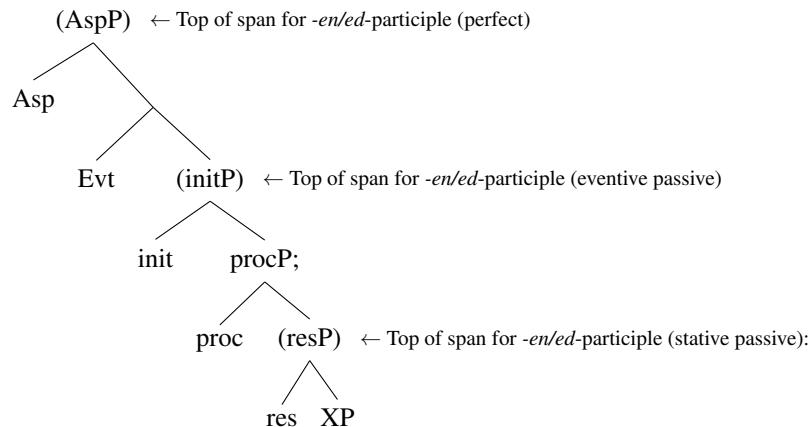


Figure 4.2

Scope of Spell-Out for the -EN/ED-Participle

The meaning of the *-en/ed*-participle turned out to be extremely weak, essentially negative, in comparison to the corresponding main verb. Its role is as the spellout of subportions of the verbal denotation. While the passive participle form was clearly within the lowest symbolic conceptual domain of the clause, the perfect participle was seen to be higher residing at the lowest point of the temporal-inflectional domain. To what extent can we see these as the ‘same’ participle, then, given all the differences we have noted between them. Compared to the specification of the uninflected root, the passive participle is consistently a ‘stunted’ version of that root. If we consider the larger inflected

verbal form, though, a stunted version of that lexical item would include the possibility of what we have assumed for the perfect participle—a version with ‘agreement’ and tense features missing. To unify the participle in English then, we could say that it is a stunted version of the inflected verbal form. To get the distinction as expressed in a language like Swedish, we could say that the passive participle is a stunted version only of the verbal root. The supine would have to be the stunted form for just the inflected forms of the corresponding verb.

This chapter has highlighted once again the difference between the conceptual event domain which traffics in verbal symbols expressing partial properties of events generalized to abstract away from time and place, and the situational domain which involves the addition of spatiotemporal information and the assertion that these event properties describe an instantiated event. I summarize the differences between these two distinct zones of linguistic structuring below.

- **Relations in the Event Essence Domain Give Rise to Pseudo-intensionality:**

Event descriptions and derived event descriptions are related to each other at an essential level, and do not entail or require mutual instantiation in the real world. We saw this with the creation of stative result participles which are semantically related to the corresponding events that have resPs, but which did not have real world ‘event’ implications. Crucially this is not to be captured in a model which builds possible worlds on top of an extensional reality, the abstract property is a basic primitive, arrived at by abstraction from experience. It has the flavour of classical intensionality but it is not.

- **Relations in the Situational Domain Give Rise to Real World Entailments:**

Situational descriptions have time and world parameters; existentially binding these variables entail actual instantiation of these situations at some time and world. With the perfect, a reference situation was introduced, where the existence of an evidential situation s' for another situation s_0 entailed the existence of that dependent situation by definition.

- **Relations found in the Event Essence Domain are in competition with lexical verb forms**

The Event domain builds basic event descriptions with all their force dynamical specificity and basic argument relations. Formatives within this domain can be sensitive to verbal subclass because they can select for domain internal structure, or undergo competition for insertion at this spell out cycle. This is the domain where elements of D_μ are first merged directly. One could think of the first phase of event building as the domain of lexical syntax in the sense of Hale and Keyser (1993a). Formatives that merge outside

this domain, as for example with the auxiliary *have*, can never be selectively blocked by verb class. Thus we found that the progressive and the passive (both eventive and stative) had verb class specific distributions, while the perfect in English applies to *all* verb classes.

• **Relations at the situational level can be modified by temporal adverbs**

The situational domain builds properties of situations that have temporal and worldly aspects. They can therefore be modified by temporal adverbials. By hypothesis, constituents in the lower event domain do not host temporal adverbs, although they do allow adverbials of manner and instrument. The Zonal hypothesis is therefore also intended to feed ordering restrictions on adverbial modification, much as in Ernst (2002).

Another important thing that has emerged from the investigation so far is the phenomenon of Blocking. Contrary to standard D(istributed) M(orphology) ideology, I have employed a general system of phrasal blocking, whereby a simple non-auxiliated verbal lexicalization always blocks the auxiliated version that spells out the same representation. Since otherwise my starting assumptions are rather different from the ones found in DM, it is not clear that the prohibition against phrasal blocking found there (see Embick and Marantz 2008) is something that I should expect to carry over to the present system. Phrasal blocking is a coherent option for the system I am assuming here where lexical verbs are specified with category features and span chunks of phrasal projections (see also Bye and Svenonius 2012, Svenonius 2012). It remains to be seen how a DM-like system would cope with the patterns and generalizations that I am capturing with phrasal blocking. To the extent that this intuition fails to translate in a similarly elegant way in DM, it would be an argument in favour of the kind of phrasal spell-out system I have been assuming. In the progressive chapter, I also assumed a blocking mechanism to prevent *-ing* applying to stative verbs, although this was blocking not under strict phrase structural identity, but more of a kind of economy condition penalizing less economical lexicalization of identical structural semantic content. I summarize the blocking facts below in (52), and state a general descriptive principle in (53).

(52) • ‘Attach *-ing* to any complete event structure and fill in with dummy verb *be*’:
blocked by stative verbs.
• ‘Spell out resP as *-en/ed* participle and spell out Evt with dummy verb *be*’
blocked by adjectives in the case of deadjectival verbs.
• Spell out procP as *-en/ed* participle and spell out Evt with dummy

verb *be'*:
blocked by unaccusative verbs.

(53) **Blocking of Auxiliation:**

In cases where a single verbal lexical item generates the same Event description as an Auxiliary structure, expression by means of an auxiliary is blocked.

The principle in (53) together with the unified denotations I have assumed for *-ing* and the *-en/ed*-participle respectively allow us to avoid any specific constraint against applying the progressive to stative verbs, or against applying passive to unaccusatives, or against forming stative passives from verbs like *empty* and *open*.

One major difference between the *en/ed*-participle and the *-ing*-participle in English has been in whether the participle involves the Merge of an additional head or not. In the case of *-ing* I argued that a head corresponding to the spell-out */iN/* is merged and has a specific semantic effect of constructing a derived identifying state. The *-ing* participle is then built by the spell out of the bare root followed by the suffix *-ing*. In contrast to this, the *en/ed*-participle cannot be decomposed in this way—it is the spellout of a span of structure, one that is systematically related to the span spelled out by the corresponding main verb in lacking uninterpretable T features (much as in the original Kratzer intuition). Moreover, in order to capture the unity behind the *-en/ed*-participle, I argued that it is associated with all subchunks of the syntactic and semantic specification of the root verb. The bare uninflected verbal form in English also differs from the *-en/ed* participle in that the latter must be morphologically specified to prohibit further suffixation.⁴⁷

The syntactic specifications for the lexical items used so far are repeated here in the list below.

(54) (a) Inflected Transitive Verb: $\langle \text{Asp} \text{ (plus uT)}, \text{Evt}, \text{Init}, \text{Proc}, \text{Res} \rangle$
 (b) Bare Root Form: $\langle \text{Evt}, \text{Init}, \text{Proc}, \text{Res} \rangle$
 (c) Participle in *en/ed*: $\langle \langle \langle \text{Asp without uT} \rangle \text{Evt}, \text{Init}, \rangle \text{Proc}, \rangle \text{Res} \rangle$
 (d) Dummy *Be*: $\langle \text{T}, \text{Asp}, \text{Evt} \rangle$
 (e) Perfect *Have*: $\langle \text{T}, \text{Asp} \rangle$

⁴⁷ While I have implemented the new architecture using a specific set of assumptions and tools, the purpose of this monograph is not to argue explicitly for a particular toolbox. My main concern is the ontological revisions proposed and the morphological and spell-out implementations are offered for explicitness and to demonstrate proof of concept. I leave it open that other implementations of the new ontological commitments are possible.

Finally, the denotation for HAVE is also repeated, in (55).

(55) $[[\text{have}]] = \lambda Q \lambda x \lambda f' \lambda d \exists s' \exists f [Q(f)(d) \wedge \text{HOLDER}(s') = x \wedge f = \lambda s \lambda d [s' \text{ gives evidence for the spatio-temporal relation between } s \text{ and } d \text{ in the same world as } s'] \wedge f'(s')(d)]$

The relation of ID-STATE, built by *-ing* in the event property zone and ID-STATE_{sit} built by *have* in the situational domain are similar, but different in a way predicted by their different zonal positions: in both cases, we get stative derivatives versions that preserve identity; in the event conceptual zone the stative derivative preserves essential properties of the event concept, in the situational zone, the stative derivative gives actual evidence for the existence of the embedded event. The definitions of ID-STATE and EVID-STATE are repeated again below for convenience.

(56) **Identifying State (ID-STATE) (Definition)**

For all event descriptions P, an *Identifying-State* for P, is a stative eventuality that manifests sufficient cognitive/perceptual identifiers of the event property P.

Evidential State (EVID-STATE) (Definition)

EVID-STATE FOR $s_0 =_{\text{def}} s'$ iff s' is a stative situation (i.e. which can have a moment as its temporal parameter) which is a salient situation that provides *criterial evidence for the existence of s_0 in the same world as s'* . The existence of s' always entails the existence of s_0 .

The EVID-STATE construction effected by *have* has different consequences for situations built from different event types. This is because of the meaning postulates that regulate the temporal run time of eventualities that can be described by the aktionsartal event sorts. We used these meaning postulates to understand the asymmetrical entailment relation between the progressive and a simple verbal form (i.e. the simple verb form in the past will entail the progressive in the past, but not the other way around). These meaning postulates also did the work of deriving the different semantic properties of the perfect when attached to verbs of different aktionsart types. The meaning postulates are repeated again for convenience in (57).

(57) (after Taylor 1977)

(I). Temporal Properties of Simple Dynamic Events:

A process event must have a temporal parameter longer than a moment. If a simple process is true at an interval I, then it is true at every subinterval of that interval larger than a moment.

(II). Temporal Properties of States:

A state can have a moment as its temporal parameter. If a state is true at an interval I , then it is true at every subinterval of that interval, including at each moment.

(III) Temporal Properties of Complex Events:

An event with complex subevental structure must have temporal run times corresponding to *each* of the subevents in that structure. If a complex event is true at an interval I , then we cannot guarantee that there is any subinterval of I at which the complex event is true.

At this point, the fact that the stative proxies are constructed so liberally in English in both the zones we have looked at, is a description—a consequence of the empirical ground we have had to account for. The identifying relationship is one which builds a stative snapshot while maintaining identity. I speculate that its liberal use in English could be due to the feature of the present tense in English which requires identification with the speech moment. Following Taylor (1977), I assume that it is only stative situations that can be identified with a single utterance moment. If anchoring in English is set up so that present only works via identity, then the operation of converting extended situations to snapshot versions of themselves will be an important and pervasive device. Thus, English has the kind of present tense that it does, and also has rich auxiliation.⁴⁸

Generally, I have been building up a view of the phrase structure of the extended verbal projection in which the syntactic zones of the clause are matched in the semantic ontology with entities of different ‘sorts’. The intuition that an extended projection is rooted in a particular lexical category and is adorned with successive layers of functional structure (cf. Grimshaw 1979) is given articulation here in the idea that the first phase is the zone where symbols of the language are combined in zone 1, and then deployed to make a propositional assertion about the world. We have moved into zone 2 in the second half of this chapter, where we have discussed the perfect and its temporal characteristics. Zone 2 is different from zone 1 in that the latter combines only elements of D_μ to give elements of D_μ , while the former denotes spatio-temporal properties of events and combines with elements of D_μ to build more specified spatio-temporal properties of events.

In the next two chapters, I will look at further auxiliary verbs that operate within the situational zone. This will include the modals, both circumstantial and epistemic. The circumstantial modals interact with the aspectual zone in

⁴⁸ Tense specification and clausal anchoring will be taken up in detail in chapter 7, where tense auxiliaries and anchoring by modals will be discussed.

interesting ways, and the epistemic modals interact with tense information and general anchoring. Once again, the focus will be on English, and on giving explicit denotations for English auxiliary forms, although the nature of the analysis that emerges should reveal properties of the universal spine (cf. Wiltschko 2014) more generally.

5 Modals and the Spatiotemporal Domain

5.1 Introduction

The view being pursued in this book is that in natural language, propositions are cumulatively built up from lexicalized event properties that are essentially atemporal, subsequently elaborated with spatio-temporal specification, and then anchored to the context in order to create a proposition that is evaluable with respect to truth about the world. With the treatment of the perfect discussed in the previous chapter, we are now well into the realm of temporally specified situations. In the next two chapters, I look at the auxiliary elements in English that can never be embedded under another auxiliary, i.e. those that are always the topmost auxiliary form whenever they are present: the modals, both deontic and epistemic, and, more abstractly, ‘tense’. I will assume that these elements all have as one part of their function the job of relating the situational description being built up to the utterance context, thereby anchoring the proposition and making it evaluable for truth.

However, in the case of modals, this is not all that these elements do. I will also need to present a specific semantics for modal expressions that does justice to their interpretations with respect to our judgements of necessity and possibility. Importantly for this project which is intimately concerned with relating syntactic and semantic generalizations, I need to account for the fact that many modals in English (and many other languages cf. Nauze 2008 for examples) are in fact ambiguous between deontic and epistemic interpretations. It is further a robust fact about natural languages that epistemic elements are systematically ‘higher’ in the phrase structure representation than deontic ones (Cinque 1999, Narrog 2012 *inter alia*).

As the reader is no doubt aware, there are many (longer) research works that deal exclusively with modality and the different interesting puzzles and patterns within this domain. The purpose of this work is not to address all of those issues for that would be simply impossible. However, as in previous chapters, we are in a good position to build on some of the knowledge and

insights gained from previous work, and integrate those results as desiderata in the present framework.

Modal meanings are found in all human languages, but for semanticists and philosophers they present a special kind of challenge because they deal with meanings where we as humans make assertions about hypothetical, counterfactual, or future events and in short, reason explicitly about things that are ‘not actual’. The tool usually used to formalize reference to the non-actual is the notion of a world parameter, which allows us to toggle between the real world (annotated here as w^*) and other ‘possible’ but non-actual worlds when expressing meanings. Possible world semantics as formalized by Kripke (Kripke 1959, Kripke 1963), Lewis (Lewis 1973, Lewis 1986) and later modified by Angelika Kratzer for a formal semantics for natural language modals (Kratzer 1977, Kratzer 1981, Kratzer 1991), is an important backdrop to all formal work on modality. I will be helping myself to world variables in my own formal analysis, but they will be integrated in a rather different way.⁴⁹

The purpose of the chapter is to provide a starting point for a treatment of tense and modality that is consistent with the new ontology and framework being proposed in this monograph. As with the previous chapters, this will necessitate a shift in thinking. This might seem unnecessary to those satisfied with our present theories of modality, but it is crucial if the agenda being pursued in this monograph is going to have any overall viability. The task before us therefore is to present a rethink of modal semantics that is consistent with the picture we have built up so far, but which does justice to the gains and insights already made in this domain within the classical model (essentially Kratzer 1977, Kratzer 1981, Kratzer 1991). In some cases, I will argue that the rethink simply packages the domain differently and makes some different choices with respect to the division of labour between syntax-semantics and pragmatics, but is intertranslatable in terms of the things it can express and account for. In other cases, the rethink will give us new purchase on old puzzles, make some things easier and some things harder. This will be the most interesting aspect of the enterprise. Regardless, the rethink itself is imposed from above by the demands of the new compositional framework, and constraints regarding the interface which I have argued to be independently desirable, so I will simply pursue it to see where it takes us.

⁴⁹ I eschewed possible worlds in the analysis of event essences and the intensionality supposedly found with the progressive, since I argued that the D_μ level could manipulate event properties in ways that did not involve invoking the existence of particulars. However, with modals, we are now in the higher syntactic domain where we are dealing with event particulars (also referred to as ‘situations’) where it is natural to appeal to a world parameter. (The only exception will be the case of dynamic modality, as discussed in section XX).

The plan for this chapter is as follows. First, in section 5.2 I lay out the basics of modal interpretation from the descriptive and syntactic literature, aiming for a fairly theory neutral description of what needs to be accounted for in any new treatment. A separate section on the interaction of modal interpretation with negation (section 5.3) comes next as an important set of facts that will also be taken into account in the proposed analysis. In 5.4 I give an overview of the most important semantic properties in the modal domain and state generalizations in a relatively theory-neutral fashion. Once we get to section 5.5, I will propose a formal analysis of circumstantial modality designed to fit in with the framework of this monograph. It will be shown that circumstantial modality and the perfect auxiliary essentially interleave in this inflectional domain, and that their effect on the situational description is actually rather similar semantically. They both involve the introduction of a reference situation (something that has been called the ‘perspectival situation’ in modal semantics) therefore establishing an indirect (though systematic) temporal relationship between the event situation and the context.

5.2 What We Know: Syntax

In the literature, a number of different types of modality have been distinguished. In all cases, modal auxiliaries modulate a situational description so that it does not assert the existence of that situation in the world but rather states its *potential* or likelihood for existence, given certain constraints. The two flavours of potentiality encoded by natural language modals appear to be virtual certainty or necessity (as in English *must*) on the one hand and mere possibility on the other (as in *might*, *can*, *could*).

Deontic modality involves notions of obligation and permission. Some state of affairs is asserted to be ‘required’ or ‘permitted’ by virtue of some regulations or imposed will of another, often left implicit. The subject of deontic sentences is often the Actor whose potential for action is being constrained in this way. In the example in (1-a), the Subject is being obligated to go to the party, in (1-b) the Subject is permitted to go to the party. However, this is not always the case, in (1-c) and (1-d), the Actor obligated to bring about the eventuality described is left implicit, and could possibly be the hearer.

| | | |
|-----|---|------------|
| (1) | (a) John must go to the party. | OBLIGATION |
| | (b) John may go to the party. | PERMISSION |
| | (c) The books must all be on the shelves by noon. | OBLIGATION |
| | (d) Flip flops may be worn in this restaurant. | PERMISSION |

In Dynamic modality, the hypothetical possibilities for the subject are constrained by their own internal characteristics. Here it is *always* the subject that is involved.⁵⁰

(2) (a) John will cheat at Monopoly if he gets the chance. **DISPOSITION**
(b) John can swim. **ABILITY**

We can thus say that in the case of dynamic modality we are dealing with *Des* potentials, rather than circumstantial potentials.

Epistemic modality involves potentials that arise because of incomplete knowledge on the part of the speaker. Here we see in (3-a) that the speaker thinks that by virtue of the other things she knows, she can *infer* the truth of the situation described, although she doesn't know it directly; in (3-b) the speaker assesses that its truth is merely possible given what she knows.

(3) (a) John must be in his office now. PISTEMIC NECESSITY
 (b) John might be in his office now. PISTEMIC POSSIBILITY

What do all of these modal uses have in common? Basically, as my descriptions have indicated, what they have in common is that they assert potentials rather than ‘facts of the matter’.

The different modals in English and in particular the epistemic - deontic split have been described and analysed carefully in many works going back to the sixties and seventies (Perlmutter 1971, Jackendoff 1972, Groenendijk and Stokhof 1975, Palmer 1986, Iatridou 1990, Brennan 1993, Portner 2009. The following table comes from Brennan (1993) (pg 8).

⁵⁰ Brennan (1993) also isolates a class of quantificational modality (after WHO), as given in (i) below.

(i) A snake can bite.

(‘It is sometimes the case that a snake bites’)

But I will not explicitly account for those here, but I take them to be a sub-species of dynamic modality.

Figure 5.1

Modals in English (from Brennan 1993)

| | Root | Epistemic |
|-------------|---|---|
| <i>may</i> | permission | possibility relative to speaker's knowledge |
| <i>must</i> | obligation | necessity relative to speaker's knowledge |
| <i>can</i> | (a) ability (b) permission | possibility relative to speaker's knowledge |
| <i>will</i> | disposition to behave in a certain way freq. used with future sense | necessity relative to speaker's knowledge |

In the above table, Root is a category that collapses the non-epistemic readings of a modal and can include dynamic and quantificational modal readings. Note also that in Brennan's table above, she has chosen to analyse *will* in English as a modal in its future 'tense' use. My position in this chapter will also be that there is no meaningful way of distinguishing between *will* and the traditional modals either syntactically or semantically. To this list I will add the following English modals *might*, *should*, *could* and *would*, which show moribund past tense morphology. Putting aside the morphology for a moment and concentrating on the possible readings, we can augment the table with the following forms.

Figure 5.2

Modals in English (Addendum)

| | Root | Epistemic |
|---------------|--|---|
| <i>might</i> | (permission in the past) ⁵¹ | possibility relative to speakers knowledge |
| <i>should</i> | obligation | necessity relative to speaker's knowledge |
| <i>could</i> | (a) ability in the past | possibility relative to speaker's knowledge |
| <i>would</i> | disposition in the past to behave in a certain way | hypothetical prediction relative to speaker's knowledge |

Examples demonstrating epistemic meanings for the above modals are given below.⁵²

⁵² For reasons that are currently obscure, only negative and interrogative versions of the English modal *can* allow an epistemic interpretation.

(4) *Epistemic Readings for the English Modals*

- (a) John may be tired after all that swimming, since he is not used to exercise.
- (b) John must be tired after all that swimming, since he is not used to exercise.
- (c) John can't be tired after all that swimming, since he is quite used to exercise.
- (d) John will be tired after all that swimming, since he is not used to exercise.
- (e) John might be tired after all that swimming, since he is not used to exercise.
- (f) John should be tired after all that swimming, since he is not used to exercise.
- (g) John could be tired after all that swimming, since he is not used to exercise.
- (h) (If he had really gone to the pool), John would be tired now.

In the following examples we see root modality.

(5) *Root Readings for the English Modals*

- (a) John may go to the party since he has done his homework.
- (b) John must go to the party since he gave Mary his promise.
- (c) John can go to the party since he has done his homework.
- (d) John should go to the party since he gave Mary his promise.

In the following we see dynamic modality.

(6) *Dynamic Readings for the English Modals*

- (a) John can swim.
- (b) John will swim any chance he gets.

How should we characterize the differences between these categories of modality semantically? And once we have a satisfactory semantic characterization, do the semantic types correlate in any way with the syntax. In other words, are there different syntactic representations corresponding to different subtypes?

In the subsections that follow I present and summarize what I take to be the broad modern consensus for what we know about modality both syntactically and semantically. These are the properties and generalizations that will form the basis of my own analysis in subsequent sections.

5.2.1 Thematic Relations: ‘Raising’ vs.‘Control’

Early work on the root vs. epistemic distinction argued that there was a lexical difference between the two types, with root meanings involving a subcategorization frame that included both an NP and S, and epistemic readings subcategorizing just for an S.

(7) $\text{modal}_{\text{root}}: (\text{NP}) (\text{S})$ (where NP is some NP in S)
 $\text{modal}_{\text{epistemic}}: (\text{S})$
 Jackendoff (1972)

Jackendoff argued that epistemic modals had the distribution of ‘speaker’-oriented adverbials while root modals had the distribution of ‘subject’-oriented adverbials. Similarly, a popular view in the literature has been that modals are a species of main verb and that root modals are transitive (control) predicates while epistemic modals are intransitive (raising) predicates (Ross 1969, Perlmutter 1971, Huddleston 1974). The raising vs. control analysis of epistemic vs. root modality is grounded in the intuition that in the case of the latter, there is an actual thematic role assigned to the subject, whereas in epistemic sentences there is not. This can be seen clearly in the possibility of expletive subjects in epistemic sentences as shown in (8), where the (a) sentence and the (b) sentence do not seem to differ in meaning.

(8) (a) John must be really tired right now.
 (b) It must be the case that John is really tired right now.

On the other hand, the ability reading of *can* in (9-a) is paraphrasable by (9-b) but crucially not (9-c).

(9) (a) John can swim.
 (b) John is capable of swimming.
 (c) *It is capable that John swims.

When it comes to deontic readings however, the situation is more subtle. Certainly, on the most common usage of these modals, there is a failure of paraphrase akin to the dynamic case above, indicating a definitive ‘theta role’ for the subject. In this case, in (10-a), the subject is the one who is obligated or permitted to perform the act in question. The paraphrase in (10-c) is ungrammatical.

(10) (a) Graduate students must pay the extra fees themselves.
 (b) Graduate students are obliged to pay the extra fees themselves.
 (c) *It is obliged that graduate students pay the extra fees themselves.

However, as mentioned before, there are many ‘deontic’ uses of modals where the subject is *not* the obliged (or ‘permitted’) Actor, but where the situation as a whole is the thing that is allowed or forced, and where the person who is to bring it about is left contextually implicit. In such cases, it is perfectly good to paraphrase the deontic modal with a construction that contains an expletive subject for the modal adjective (11).

- (11) (a) The shirts must be clean and colourful.
 (b) It is required that the shirts be clean and colourful.
- (12) (a) The exam may be handed in before the deadline.
 (b) It is allowed that the exam be handed in before the deadline.

Thus, while it seems right to analyse epistemic readings as taking a single sentential argument syntactically, and dynamic readings as expressing a relationship between a subject and a VP predicate, deontic interpretations seem to be found in both kinds of syntactic subcategorization pattern. While the deontic interpretations always do require some sort of Actor who is being obliged or permitted, it seems that the syntactic expression of that individual is not obligatory.

Brennan (1993) gives examples of both epistemic modals and deontic modals with expletive subjects, demonstrating that they, unlike dynamic modals, do *not* obligatorily select for an external argument.

- (13) *Epistemic*
 - (a) It may be raining.
 - (b) There may be some eggs in the refrigerator.
 - (c) It must be obvious that I have dyed my hair.
 - (d) There must be somebody drumming on the roof of the car.
- (14) *Deontic*
 - (a) There may be up to five cars in the lot at one time. (b) It must be quiet in the reading room at all times. (c) There must be three life guards on duty.

The intermediate position of deontic modality with respect to this diagnostic is paralleled by the subject scope diagnostic described below.

5.2.2 Scope wrt to a Quantified Subject

Brennan (1993) gives data from the relative scope of a quantified subject and the modal force which shows the same pattern as the paraphrase test above. Specifically, when it comes to epistemic readings, a modal can always take scope over the quantified subject, whereas in the dynamic readings it may not.

We can see this by inspecting the following sentences from Brennan (1993). In (15), under an epistemic reading, we find both a contradictory reading and a non-contradictory one.

(15) Every radio may get Chicago stations and no radio may get Chicago stations.

Contradictory Reading: $\forall x \Diamond [\text{radio}'(x) \rightarrow \text{get-Chicago-stations}(x)]$

$\wedge \neg \exists x \Diamond [\text{radio}'(x) \wedge \text{get-Chicago-stations}(x)]$

Non-contradictory Reading: $\Diamond [\forall x [\text{radio}'(x) \rightarrow \text{get-Chicago-stations}(x)]$

$\wedge \Diamond [\neg \exists x [\text{radio}'(x) \wedge \text{get-Chicago-stations}(x)]]$

On non-contradictory reading, the modal takes wide scope over the quantified subjects. If we use a dynamic modal, the ability modal *can* instead, we can see in (16) that the sentence only has the contradictory reading.

(16) Every radio can get Chicago stations and no radio can get Chicago stations.

Once we turn to the deontic modals, we see that both contradictory and non contradictory readings are possible, as in the epistemic case. (For some reason, Brennan 1993 does not give an example of this case. I have provided my own.)

(17) Every child can be invited and no child can be invited, but you are not allowed to play favourites!

See also Bobaljik and Wurmbrand (1999) for detailed argumentation that both deontic and epistemic modals need to be seen as raising predicates.

5.2.3 Modals and Symmetric Predicates

Finally, Brennan (1993) gives an interesting argument from symmetric predicates which once again indicates that the dynamic modals clearly modulate the event concept and affect the thematic role for the external argument, while epistemic modality does not. A symmetric, or reflexive predicate *R* is one for which *R*(*x*,*y*) entails *R*(*y*,*x*). Some examples in English are *shake hands with* or *get the same score as*. So for example, (18-a) entails (18-b) and (19-a) entails (19-b).

(18) (a) The governor shook hands with all the prisoners.
 (b) All the prisoners shook hands with the governor.

(19) (a) Peter got the same score as Joan.
 (b) John got the same score as Peter.

If we add modal auxiliaries to these sentences, we see that the entailment is clearly disturbed under dynamic interpretations. In (20), the (a) sentence does *not* entail the (b) sentence. Peter might have the ability to get the same score as Joan no matter how high she scores, but the same might not be true of Joan.

(20) (a) Peter can get the same score as Joan.
 (b) Joan can get the same score as Peter.

On the epistemic reading of the modal *might*, the entailments of symmetric predicates remain completely undisturbed however, as we can see in (21) where the (a) and (b) sentences are mutually entailing.

(21) (a) The governor might shake hands with all the prisoners.
 (b) All the prisoners might shake hands with the governor.

When it comes to the deontic readings, in the cases where the subject is also the Actor whose actions are being manipulated by permissions and prohibitions, the entailment is also clearly disturbed. In (22), I may have given the governor permission to shake hands with whomever he likes, but not allow the prisoners themselves to make any such overtures so that the (a) sentence does not entail the (b) sentence.

(22) (a) The governor may shake hands with all the prisoners.
 (b) All the prisoners may shake hands with the governor.

However, if we construct a deontic sentence where locus of the permission is not explicitly represented in the sentence as in (23), then it seems as if the entailment does in fact go through. If for example, you are told that in the interests of keeping the peace, you must make sure that the following comes to pass, then the version in (23-a) is exactly the same prescription as the version in (23-b).

(23) (a) The girls must get the same score as the boys.
 (b) The boys must get the same score as the girls.

Thus, once again, when the subject is the locus of permission or prohibition, the deontic modals pattern with the dynamic, and with the epistemic modals otherwise.

The situation and the diagnostics are summarized in the following table.

| | Semantic Role for Subject | Modal Scope wrt Subject | Persistence of Symmetry |
|------------------|---------------------------|-------------------------|-------------------------|
| <i>Dynamic</i> | YES | Modal Low | NO |
| <i>Deontic</i> | YES/NO | Ambiguous | NO/YES |
| <i>Epistemic</i> | NO | Ambiguous | YES |

5.2.4 Ordering and Typology

With respect to modal ordering in English, we have already noted that English has a special requirement on its modals which means that regardless of its interpretation, a modal always ends up in the highest position of the clause it appears in. Thus, it is not easy to see from the surface position of an English modal that there is actually a more subtle set of ordering patterns that have fairly robust typological support.

If we look at a language like Swedish, which is quite similar to English in many respects, but which has modals that can occur in nonfinite forms, then a more fine grained pattern emerges.⁵³

The first thing to observe is that the dynamic meaning of the modal ‘can’ in Swedish referring to ability, can appear inflected directly for past tense, as in (24).

(24) Han kunde skriva klart sin uppsats. *Swedish*
 He can-PAST write finished/ready his article.
 ‘He was able to finish his article’ (PAST > Dynamic Modal)

When we look at the interaction of the dynamic modal with the perfect auxiliary ‘have’, we find that it can be successfully embedded under ‘have’ to give (25). The other order shown in (26), is simply not possible.

(25) Han har kunnat skriva klart sin uppsats.
 He has can-PAST PART. write finished/ready his article
 ‘He was able to finish his article’ (HAVE > Dynamic Modal)

(26) Han kan ha skrivit klart sin uppsats.
 He can have written finished/done his article.
 ‘*it is capable that he has finished his article’ (*Dynamic Modal > HAVE)

Turning now to epistemic versions of ‘can’, we see the opposite ordering with respect to the ‘have’ auxiliary. It is perfectly possible to embed ‘have’ under

⁵³ The data from this section come from my own fieldwork with native speakers.

the Epistemic modal, as shown in the grammatical interpretation for the above sentence (repeated here as (27)).

(27) Han kan ha skrivit klart sin uppsats.
 He can have written finished/done his article.
 ‘It is possible that he has finished his article’ (Epistemic Modal > HAVE)

Turning finally to deontic modality, the modal auxiliary ‘must’ can get a deontic reading of obligation, and under this reading it is possible for it to embed the perfect auxiliary ‘have’ as in (28).⁵⁴

(28) Han måste ha gjort leksan innan Fredag.
 He must-PRES have done the homework within Friday
 ‘He must have the homework done by Friday.’ (Deontic Modal > HAVE)

Unfortunately, in Swedish, ‘must’ is one of those modals that has no non-finite forms (like the English modals), so to check if deontic modality can be embedded under ‘have’, we need to choose another more inflectionally flexible deontic modal. The clearest possibility deontic modal in Swedish here is *få*, which is the most common way to express permission. The most natural ordering for this modal and the perfect auxiliary is with the perfect scoping over the modal as in (29).

(29) Han har fått göra det hele sitt liv.
 He has got-PAST PART. do that whole his life
 ‘He has been allowed to do that his whole life.’ (HAVE > Deontic Modal)

Finally, when two modals cooccur, the epistemic modal is always higher than the root modal.

(30) Han måtte kunnat skriva klart sin uppsats
 He must-PAST can-PASTPART write-INF his article
 ‘He must have been able to finish writing his article.’

⁵⁴ Note that here, the reading is that the perfect eventuality is obliged to hold some time in the future. This is consistent with deontic modality in general, which is obligatorily forward shifting with respect to the evaluation time. These sentences are tricky to construct felicitous versions of because it requires some context to construct a situation where a perfect state is going to be relevant in the future.

So we can see from this data that ‘have’ is higher than Dynamic modality and lower than Epistemic modality, but that it seems to be in principle interleavable with Deontic modality. I will assume that this is true in English also, and place the base position for deontic meanings within the same sortal zone as the perfect auxiliary discussed in the previous chapter.

This pattern is found quite generally crosslinguistically (see e.g. Nauze 2008). Nauze (2008) finds broad support for the following hierarchy proposed also in Cinque (1999).

(31) Epistemic modality < Circumstantial Modality < Dynamic Modality

In English however, word order never gives us any direct evidence for this hierarchical order, although the diagnostics above concerning the subject position and scope are indirect evidence for the very same ranking

Another suggestion of base position ordering even in English comes from the taxonomy of modals themselves which seem to be in an interesting implicational relation with respect to the readings they allow. If we list the three types of reading, dynamic, deontic and epistemic according to the height established via crosslinguistic typology, we can see that if a modal has a ‘lower’ reading in English then it also has all the corresponding higher readings, but the reverse implication does not hold.

| | Dynamic | Circumstantial | Epistemic |
|------|-------------------|-------------------|--------------------|
| | <i>can/could</i> | <i>can/could</i> | <i>can't/could</i> |
| | <i>will/would</i> | <i>will/would</i> | <i>will/would</i> |
| (32) | | <i>may</i> | <i>may</i> |
| | | <i>must</i> | <i>must</i> |
| | | <i>should</i> | <i>should</i> |
| | | | <i>might</i> |

In terms of closeness to the root, it should also be noted that the suffix *-able* which creates adjectives with a ‘modal’ flavour from verbal roots in English, is confined to dynamic meanings, the lowest of the three possibilities, where the internal properties of the subject are what are at stake.

(33) (a) The book can be read (circumstantial or dynamic).
 (b) The book is readable (dynamic: inherent property).

The same seems to be true of the suffix *-er*, which creates Nouns with a generic/type sense with the flavour of dynamic dispositional modality.

(34) (a) John will cheat at Monopoly. (dispositional or futurate)
 (b) John is a cheater. (dispositional only)

5.2.5 Taking Stock

The view from the syntactic literature thus suggests that the different classes of modal—dynamic, deontic and epistemic—correlate systematically with different syntactic behaviours which all implicate height of attachment. Firstly, there seems to be a major zonal rift between dynamic modality on the one hand and the other two categories. In Figure 5.3 one can see that the properties of dynamic modality strongly implicate the lowest domain of nonspatio-temporal eventive properties as expressed by elements of D_μ . Dynamic modulations are difficult in principle to separate from the event description itself, they show no scopal ambiguity with respect to higher elements, and they can be expressed by derivational suffixes such as *-able* and *-er*.

Figure 5.3
 Distinguishing Dynamic Modality

| Dynamic Participant Internal Modality | Circumstantial/Epistemic Participant External Modality |
|---|---|
| Expletive subjects impossible | Expletive subjects possible |
| Modalizing affects the symmetry of the predicate | Variable |
| Occurs lowest in a sequence of modals crosslinguistically | Variable |
| Scopes low wrt to (polar) negation | Variable |
| Scopes low wrt to Subject position | Variable |
| Can be expressed by derivational suffixation in English | No |

With respect to circumstantial vs. epistemic, we can motivate a syntactic height distinction here as well internal to this higher category. The differences I will assume to hold between the two higher categories are shown below in the table in Figure 5.4.

Figure 5.4
 Distinguishing Circumstantial from Epistemic

| Circumstantial Modality | Epistemic Modality |
|--|---|
| lower than epistemics under cooccurrence | higher than circumstantial under cooccurrence |
| Morphological PAST can shift modal state in some lgs can take scope under the subject | modal state cannot be shifted by morphological PAST never take scope under the subject |
| No interaction with Speaker oriented meanings | Speaker oriented |

It is still important to keep in mind however that the modals in English actually show no difference in *surface* height. The following facts hold for *all*

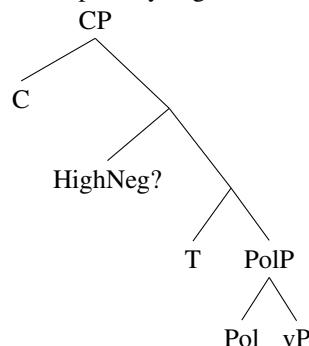
the English modals: a modal never co-occurs with any other modal; they are always the highest auxiliary when they occur; they invert in questions and can form tags; they can host clitic negation.

(35) (a) *John must can do that.
 (b) John must be teasing the cat.
 (c) Must John always tease the cat?
 (d) John mustn't tease the cat.

This does not mean that modal interpretations cannot be correlated with height of attachment in English. It means only that *whatever the initial Merge position* of a modal is, in English it obligatorily raises to the position that used to be called INFL. We will discuss the role of modals as anchoring elements in the sense of Wiltschko (2014) in chapter 6.

5.3 Interaction with Negation

In this section, I examine the scope of modal meaning with respect to the scope of negation in English. This is an area that could potentially give us semantic evidence for the base position of different modal elements, if a clear pattern could be found. The first problem we encounter in evaluating the data however will be that the position of negation itself is less clear and reliable from a typological point of view than the other verbal elements discussed so far in this monograph. This is compounded by the fact that different types of negation have been distinguished: constituent negation and metalinguistic or 'high' negation, in addition to clausal polarity negation. Zanuttini (1992), Zanuttini (1997) argues for at least two possible positions for negation in Italian, even when just polarity negation is taken into account. Schwarz and Bhatt (2006) argue for 'high negation' which has different presuppositional properties from standard polarity negation.



Care must therefore be taken when examining the data, that we keep the type of negation under consideration fixed. First, I summarize the facts of English modals with respect to scope of negation searching for a generalization, and then I explore a purely semantic account of the patterns.

5.3.1 English Modals are Idiosyncratic

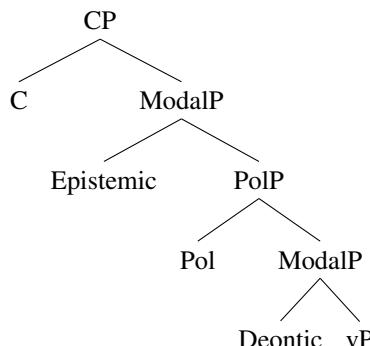
Interestingly, most modals in English seem to be categorical in their preference for a scope position with respect to polarity negation regardless of their flavour in any particular usage. (see Cormack and Smith 2002).

(I) Modals that scope over polarity negation: *should, must, will, would, might, may* (*epistemic*)

(II) Modals that scope under polarity negation: *dare, need, can, could, may* (*deontic*).

If we try to range these modals with respect to polarity, the natural option would be the one in (36)

(36)



However, this picture is immediately falsified by *should* in its epistemic and deontic uses, which consistently scope above negation.

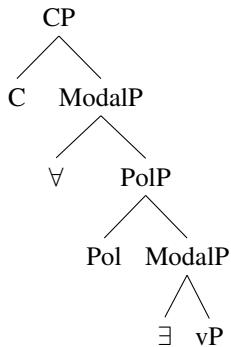
(37) (a) John shouldn't eat so much chocolate. SHOULD>¬
 (b) John shouldn't miss his plane—he left early. SHOULD>¬

And by *could* in its epistemic and circumstantial uses, both of which scope under negation.

(38) (a) John couldn't go to the party. ¬>COULD
 (b) John couldn't be there yet—he left way too late. ¬>COULD

One might think that perhaps the ordering with respect to negation is more sensitive to the modal force of the modal in question, and the hierarchies are more semantic in nature. A first try would be the ordering shown in (39).

(39)



However, this doesn't work either because we find universal modals that scope under negation, as well as existential modals that scope above it, as shown in (40).

(40) (a) John doesn't need to go to the party. $\neg > \text{NEED } (\forall)$
 (b) John might not go to the party. $\text{MIGHT } (\exists) > \neg$

To repeat, even when a modal has multiple readings (dynamic, deontic, epistemic) the scope with respect to negation carries over to all uses of that modal. This suggests that these ‘ambiguous’ modals are a single lexical item, and that their requirements with respect to negation are an integral part of their meaning, regardless of ‘height’ of attachment.⁵⁵

But maybe English is especially rigid not because of something to do with modal meanings per se, but because the English modals all have a uniform final position in the sentence and none of them possess nonfinite morphological forms. While it is unclear what effect this might have, we can take a shortcut in our search for a linguistic generalization by comparing English with its closely related Germanic cousin, Swedish. Swedish has the modal *måste* which is like English in having both deontic and epistemic interpretations, and furthermore like English in having no non-finite forms.

⁵⁵ The only exception to this generalization is the modal *may* which does have different scopes with respect to negation depending on epistemic vs. deontic interpretation. I will assume that this means that at this point in the history of English, these have become two distinct lexical items

(41) Nigel must not do that.

(42) Fredrik måste inte göra det
 Fredrik must not do-INF that
 'Frederick doesn't have to do that' ($\neg > \forall$)

A comparison of the two sentences in (42) reveals that Swedish forces an interpretation where negation scopes above the modal *måste*, whereas in English as we have seen negation scopes consistently below *must*.

Micro-Comparison I

| Etymon | Force | Modal Base | Scope wrt Neg | Non Finites |
|--------------|-----------|-------------------|---------------|-------------|
| <i>måste</i> | \forall | Deontic/Epistemic | Below | No |
| <i>must</i> | \forall | Deontic/Epistemic | Above | No |

Comparing now the English *can* with Swedish *kunna*, we find that although the Swedish modal shows a great variety of nonfinite forms and the freedom to occur under the perfect and under other modals, the scope with respect to negation is identical to that of English. See the examples in (43).

(43) Nigel can't do that.

(44) Fredrik kan inte göra det
 Fredrik can not do-INF that
 'Fredrik can't do that' ($\neg > \exists$)

This second microcomparison is also summarized below.

Micro-Comparison II

| Etymon | Force | Modal Base | Scope wrt Neg | Non Finites |
|--------------|-----------|-------------------|---------------|-------------|
| <i>kunna</i> | \exists | Deontic/Epistemic | Below | Yes |
| <i>can</i> | \exists | Deontic/Epistemic | Below | No |

It seems as if there is no hope for a generalization based on the simple parameters of modal meaning, or surface syntactic positioning. There are two sort-of generalizations that emerge from the data, although they are not easy to encode in a simple hierarchy, those generalizations are given in (45).

(45) (a) **Interpretational Constraint *ALLOW-NOT**: There are no deontic, existential interpretations that scope over negation (Iatridou and Zeilstra 2013 also say they found none such).
 (b) **Lexical Constraint *NOT-NECESSARY**: There are no universal modals that are *just* epistemic that scope under negation.

These generalizations have a slightly different status from each other since the first is a constraint on interpretations, while the second is a constraint on modal lexical items.

I conclude that the failure to find a clean correlation here tells us something important about modal syntax and semantics. It looks like the data above force us into an analysis where each modal *selects* for negation in its own idiosyncratic way. In what follows, I am going to propose that this is a form of *syntactic selection* (c-selection) which is not reducible to an independently establishable semantic property, although it has semantic implications.

5.3.2 A Semantic Account of the Idiosyncrasy?

The puzzle of the interaction of modal quantificational meanings which negation has of course received attention in the semantic literature Iatridou and Zeilstra 2013, Israel 1996, Homer 2012). The consensus from the aforesaid works is that modal meanings come with a specific ‘polarity’ sensitivity.

Here, I summarize in a nutshell the proposal of Iatridou and Zeilstra (2013).

- Some modals are PPIs, some are NPIs and some are neutral.
- All modals in English (regardless of flavour) are generated in a position below I and indeed below Neg.
- All modals in English (regardless of flavour) move to I because they are tensed.
- This movement reconstructs optionally depending on the polarity requirements of the particular modal.
- The reconstructed position is the default, and is the only one available for the neutral modals.
- The PPI modals are not allowed to reconstruct because of their polarity sensitivity and hence only get interpreted in the ‘high’ position

Iatridou and Zeilstra only look specifically at deontic modality, and their classification of the modals in English with respect to polarity is given in the table below.

| | PPIs | Neutral | NPIs |
|-------------|-------------------------------|-------------------------|-------------|
| Universal | <i>must, should, ought to</i> | <i>have to, need to</i> | <i>need</i> |
| Existential | | <i>can, may</i> | |

The analysis in Homer (2012) is similar to Iatridou and Zeilstra (2013) in that the base position of all modals is below PolP, and that it is the positive polarity modals like *must* that are special. For Homer, a modal like *must* has to move out to take scope over negation in order to capture the PPI effects.

However, the semantic analysis of PPI-ness is truly only an *independent* explanation from the semantic side, if we can show that there is something about the semantics of these items that gives rise to their polarity behaviour. In fact, formal semanticists know a lot about the meanings of PPIs and NPIs and the reasons why they behave the way they do. Unfortunately, even though there are some distributional behaviour facts in common with traditional NPI and PPI behaviour, many of the explanations and decompositions do not seem to be independently verifiable for ‘modal’ PPIs. (See Homer (2012) and Iatridou and Zeilstra (2013)). I quote from the latter here:

"How can we prove that *must* (and the other PPI-modals) obligatorily introduce domain alternatives? In truth, we cannot. The reality is that the diagnostics to detect whether some element introduces domain alternatives in general are very weak. . . . The only thing that can be said is that if elements like *must* are assigned the ability to introduce alternatives, then it may be possible to express their PPI-hood in certain pre-existing terms. Moreover, if such an analysis is correct, the question arises as to why *must* is only banned from anti-additive contexts and not from all downward entailing contexts. Another problematic question that arises is the following: since nothing specific in the discussion of polarity hinges on the choice of world variables instead of variables over individuals, we would expect PPIs and NPIs to have either universal or existential force in both. This is not what is found, however. In the domain of individuals, PPIs and NPIs only seem to come with existential force; no polarity items with universal force over individuals seem to exist. However, in the modal domain, PPIs and NPIs are found on items with universal force, and at least among the deontic modals, no PPIs or NPIs with existential force have been reported." (Iatridou and Zeilstra 2013)

It is also notable that Homer at least is forced to stipulate an overt scopal movement for *must*. But this should be seen as quite striking and unusual given all that we know about quantificational elements and their scopal interactions. With respect to adverbial elements, and negation in particular, the facts seem to be that such elements are interpreted in their ‘surface? structure position, assumed in these cases to be the same as the base generated position. With respect to quantificational DPs, they can often appear displaced from what is assumed to be their base position, and they also seem to give rise to ambiguity in their scopal relationship with other items. So if modals are quantificational elements, do they pattern with negation or do they pattern with quantificational DPs in their scope taking behaviour? First of all it seems as if the English modals do not surface in their base generated position, not if we are going to tie differences in the modal base to height of first merge. But they are not interpreted uniformly with respect to their base positions, and neither are they interpreted uniformly with respect to their surface positions. Specifically for the semantic proposals discussed above, we could ask: Why don’t all modals scope under negation if they are base generated there? Or similarly:

Why don't they all scope over negation if they all end up higher than polarity negation on the surface? Somehow we have to say that modals like *must* do not reconstruct under polarity negation for scope, while modals like *can* obligatorily reconstruct. What is the independent *semantic* fact that guarantees this? Especially given the fact that English *must* and Swedish *måste* are otherwise identical in range of interpretations, but differ with respect to PPI-ness. Further, why should head movement have an effect on scope, and then only for certain lexical items?

(46) The Difference between English and Swedish MUST under a Reconstruction Account:

Swedish: *måste* obligatorily reconstructs for scope wrt polarity negation.

English: *must* cannot reconstruct for scope but must take scope in its surface position.

If we allow more flexibility of ordering within the second phase domain, and allow in particular free attachment of Polarity Phrase (PolP) within that domain, then there is an alternative way of stating the facts. Basically, the ordering of circumstantial modality and PolP is in principle free in this system, and the actual order found for each modal depends idiosyncratically on selection. In this version, the behaviour with respect to negation is due to syntactic selection under locality, with no movement for scope taking allowed or required.

(47) The Difference between English and Swedish MUST under a Selection Account:

Swedish: *måste* selects for a polarity-neutral property of situational properties.

English: *must* selects for PolP within the situational domain.

If we assume that *may* is actually two items, then *may*₁ (deontic) and *may*₂ (epistemic), then the English modals have the following selectional properties.

Box 5.1
English Modals and Selection for PolP

| SELECTION | |
|---------------------------|--|
| Polarity Phrase | <i>should, must, will, would, might</i> <i>may</i> ₂ |
| Situational Phrase | <i>can, could, need</i> <i>may</i> ₁ |

Comparing this directly to the PPI account, it the so-called PPI modals come out as those that *select for* PolP. Let us see how this selectional restatement ends up giving the correct effects, together with other plausible assumptions about the functional sequence.

Assumptions:

- High Negation sits in the left periphery (the extended C-domain); low negation (constituent negation of verbs) is within the first phase; PolP is in the inflectional, second phase zone of the clause and is the label I give to traditional clausal/polarity negation.
- There can be only one PolP within the second phase. If it is set as positive, there is no overt morpheme in English. Every inflectional domain/2nd phase must contain a PolP.
- PolP has no rigid place within the second phase.
- Interpretation of any element with respect to negation is based on the first merge position.

If a modal selects for PolP, it means that you will not see this overtly when the polarity is positive, but it will scope above any expressed polarity negation. Modals that select for a polarity neutral situational description, will always have PolP merged above them at some stage and this will mean that polarity negation will always scope over those modals. Thus, when Pol is negative, it is always interpreted below *must*, and above *can*. The classification is lexical item by lexical item, not reading by reading, as would be expected if this were a lexical c-selection property. It also accounts for the locality of the effect and for the fact that it concerns polarity negation specifically, and not 'high' negation.

The more explanatory question that we posed to the PPI account also needs to be addressed to the selectional account. If this really is a matter of syntactic selection, *why* should it be that there is such a close selectional relationship between modal meanings and negation, and how does it affect the compositional semantics of these forms? We will address this question in section 5.5, but for

now we merely note that circumstantial modality must interleave with PolP in the inflectional domain, and whether a particular modal combines with a polarity specified projection or not seems to be a matter of lexical specification at the level of c-selection.

5.4 Semantic Properties of Modals

5.4.1 The Classical Theory

Up to this point, I have deliberately refrained from giving an introduction to the classical model for interpreting modal meanings in natural language, the Kratzerian model (Kratzer 1977, Kratzer 1981, Kratzer 1991). There are many useful such introductions in the literature which I refer the reader to for a more step by step and technically explicit exposition (Portner 2009). I offer a more schematic introduction here concentrating on the intuitions and logical underpinnings of the system. My own analysis will not require an engagement with the technical details of the classical model *per se*, but it is important to see how it relates to and differs from it.

For an understanding of the systems of logic underpinning modal reasoning, I am indebted to McNamara (2014), which has inspired some of the initial presentation, and to whom I refer the reader for more detailed references.

The formal tradition for thinking about these things comes from the philosophers and the logicians. They were concerned with logical relations among propositions first and foremost and not with the inner workings of natural language, but in the realm of epistemic modality there are clear natural language expressions that map nicely to these notions. Possibility and necessity (conventionally notated as \diamond and \Box respectively) can be thought of as a set of notions that can partition the domain of propositions into jointly exhaustive and mutually exclusive subclasses according to the schema shown in Figure 5.5.

Figure 5.5
Alethic Modality

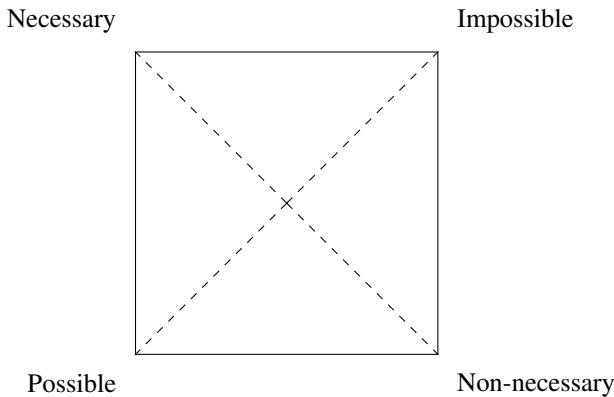
| | | |
|---------------|------------|------------|
| Possible | | |
| Necessary | Contingent | Impossible |
| | | |
| Non-Necessary | | |

If we express this in terms of a square of oppositions, we get the figure in 5.5, where the dotted lines represent the contradictions in the above square of oppositions; entailments follow the vertical lines from top to bottom; the

topmost horizontal line links contraries, and the bottommost horizontal line links subcontraries.

Figure 5.6

Alethic Modality: Square of Oppositions



The logic is usually axiomatized based on the ‘Necessary’ primitive, from which all of the other notions can be defined with the help of negation.⁵⁶ In addition it is usually assumed that the two statements in (48) hold:

(48) (i) If $\Box p$, then p
 (if a proposition is necessary then it is true), and
 (ii) If p , then $\Diamond p$
 (if a proposition is true, then it is possible),

These indeed seem like rather unremarkable and harmless assumptions.

A natural analogy allows us to move from the alethic domain to the domain of quantification. Here, necessity has its parallel in universal quantification, and possibility as existential quantification. We can see how the partition works out in this case in Figure 5.7.

⁵⁶ Necessity is classically chosen, but in fact, any one of these modal notions can be taken as basic and the others derived from it with the help of negation.

Figure 5.7

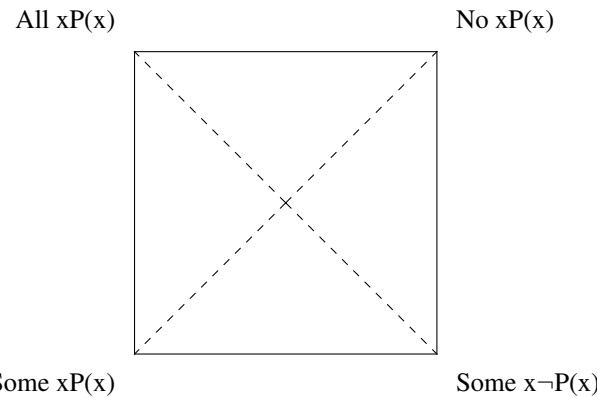
Quantification

| | | |
|-------------|---------------------------------------|-------------------|
| | Some $xP(x)$ | |
| All $xP(x)$ | Some $xP(x) \wedge$ Some $x\neg P(x)$ | No $xP(x)$ |
| | | Some $x\neg P(x)$ |

So that In terms of a square of oppositions, we get the figure shown in 5.8

Figure 5.8

Quantification



The reason I am spelling this out in perhaps overly obvious detail, is that I wish to highlight the fact that the classical model for interpreting modality in terms of possible worlds is a rewriting based on the success of the analogy between the two schemes above the quantification square and the alethic square.

I quote McNamara (2014) here who says:

“These deep quantification analogies reflect much of the inspiration behind ‘possible world semantics’ for such logics. Once the analogies are noticed, this sort of semantics seems all but inevitable.” McNamara (2014), pg

Once the analogy is noticed, It is possible to give a (Kripke-an) possible world semantics for modal logic in the following way. We assume a set of possible worlds W . We also assume a relation R relating worlds to worlds that have something in common. Formally, we can think of a model M as consisting of a pair W and R and a valuation function that assigns truth values to propositions in a particular world. (We assume that propositions are either

true or false in a world, and never both.) We can call the set of all worlds v for which the relation $R(w, v)$ holds the set of all worlds that are *accessible* from w . Intuitively, the set of all accessible worlds from the base world is going to be the set that the modal operator quantifies over.

Formally, therefore, we can translate the notion of necessity into universal quantification in the following way, with definitions taken from Portner (2009). Skipping the standard axioms relating to the general propositional calculus, we cut right to the chase and give the definitions for necessity modals and possibility modals in (49-a) and (49-b) respectively.

(49) (a) α is of the form $\Box\beta$ and for all v such that $R(w, v)$, $\llbracket \beta \rrbracket^{v, M} = 1$.
 $(\Box\beta$ is true iff β is true in *all* members of W accessible from w).
(b) α is of the form $\Diamond\beta$ and for some v such that $R(w, v)$, $\llbracket \beta \rrbracket^{v, M} = 1$.
 $(\Diamond\beta$ is true iff β is true in *some* member of W accessible from w).

The beauty and elegance of the ‘Simple Modal Logic Hypothesis’ as Portner calls it, is that “the meaning of every modal expression in natural language can be expressed in terms of only two properties: (a) whether it is a necessity or possibility modal, and (b) Its accessibility relation R .” (Portner 2009, pg 31). So for example, we can imagine an epistemic frame as in (50).

(50) **Epistemic Frame:**

$F = \langle W, R \rangle$ is an epistemic frame iff for some individual i :
• W = the set of possible worlds conceivable by humans.
• R = the relation which holds between two worlds w and w' iff everything which i knows in w is also true in w' .

We can further extend the analogy to the deontic frame, and posit a domain of quantification defined by the accessibility relation corresponding to the ‘rules’ established by a certain context. This is given in (51).

(51) **Deontic Frame:**

$F = \langle W, R \rangle$ is a deontic frame iff for some system of rules r :
• W = the set of possible worlds conceivable by humans.
• R = the relation which holds between two worlds w and w' iff all of the rules which r establishes in w are followed in w' .

In fact, there is no limit to the contextual specificity of the modal bases one could imagine. There could be subvarieties of deontic modal bases according to what kinds of laws or desiderata are involved and these could be filled in by context. We could also imagine other modal bases such as *buetic* conver-

sational backgrounds related to wishes and *teleological* conversational backgrounds related to aims.

That the nature of R is flexible and subject to both contextual and linguistic modulation by means of framing adverbials, has been amply demonstrated in Kratzer (1977). In (52) we see two examples where the adverbial clause makes explicit the nature of the R that is being assumed.

(52) (a) In view of the laws of Massachusetts, drivers must yield to pedestrians.
 (b) In view of the traditions of our family, you, as the youngest child, must read the story on Christmas eve.

Because of the variability of deontic interpretations, which range all the way from explicit commands and permissions to general circumstantial facilitating and inhibiting conditions, I will henceforth refer to this class of modal interpretations as *circumstantial* modality, essentially following Kratzer's terminology.

However, there is some reason for disquiet. If we are explicit about the analogy between the alethic and quantificational schemas and the one we would set up for permissions and obligations, we see that things do not work out quite the way one would wish. Treatments of classical deontic logic do in fact operate with a parallel picture of the partition of notions shown in 5.9. Instead of necessity we have obligation (OB(p)) and instead of possibility we have permissability (PE(p)). As with the alethic case, these modal notions are assumed to partition the space of propositions exhaustively and are mutually exclusive.

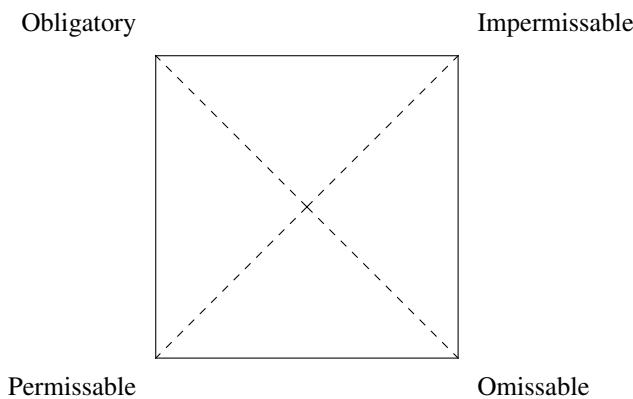
Figure 5.9
 Deontic Modality

| | | | |
|-------------|----------|---------------|--|
| Permissible | | | |
| Obligatory | Optional | Impermissible | |
| Omissible | | | |

This gives rise to the square of oppositions shown in 5.10.

Figure 5.10

Deontic Modality Square of Oppositions



This seems all very well and good except for the fact that unlike in the other two cases, the simple entailments in (52) do not seem to go through.

(53) (i) OB(p) does not entail p
(ii) p does not entail PE(p).

In other words, just because there is an obligation on someone to do something, it does not mean that it gets done. And if p is the case, it does not mean that it was permitted.

The original Kratzerian system has the advantage of modelling the unity behind the different meanings of the English modals in different instances of use, but it does this only if we choose to ignore the logical disanalogies between the quantificational system and the logic of permission and obligation. In addition, the classical system fails to connect the systematic differences between circumstantial modality and epistemic modality with respect to syntactic height in the verbal extended projection.⁵⁷ As in the classical system, we want to build in the intuitive similarities between epistemic and circumstantial flavours of modality, but quantification over different modal bases is not the

⁵⁷ Kratzer also innovates the idea of the Ordering Source on worlds, a third parameter which is necessary to account for our judgements of the gradability or degrees of possibility and/or obligation. I will not be making use of the ordering source or its equivalent. I note that the examples used to demonstrate the necessity of this move come from modal expressions that are either nominal or adjectival. I suspect that the factor of gradability comes from the adjectival domain itself and is not a core property of verbal auxiliary modals. The investigation of this idea is obviously beyond the scope of this short monograph.

only way to express those similarities and differences between the different levels of modal expression. I will retain the use of possible worlds to express hypotheticality and potential in the semantic representation, but I will employ a semantics of ‘choice’, rather than universal vs. existential quantification over possible worlds. Only in the case of logical choices, will the two schemes be equivalent.

To solve the problem of relativizing the modal interpretation to height of attachment, it is important to note that great strides have been made in this direction in recent work by Valentine Hacquard in her dissertation and many subsequent papers (Hacquard 2006 etc.). In that work, she made an important and influential proposal extending the Kratzerian system to account for generalizations at the syntax-semantics interface. In particular, Hacquard is at pains to reconcile the elegance of the Kratzerian system where a single underspecified meaning can handle both epistemic and root interpretations, with the results of linguistic typology (cartography) which suggest the generalization that epistemic readings attach higher in the clause, *outside* tense, and root meanings attach *inside* tense. She proposes a system which ties particular types of interpretation to height in the structure. Her idea is to replace the base world from which the modal base is calculated with an event instead, and relate the semantic differences to differences in how that event is anchored. The in turn is sensitive to the height of the modal in question. Specifically, she claims:

- (i) when the modal is speaker-oriented, it is keyed to the speech time and receives an epistemic interpretation;
- (ii) when the modal is attitude holder-oriented, it is keyed to the attitude time and receives an epistemic interpretation;
- (iii) when the modal is subject-oriented, it is keyed to the time provided by tense and receives a root interpretation.

Hacquard otherwise keeps intact the central structure of the Kratzerian solution: modals are functions from sets of possible worlds to sets of possible worlds; a restriction via contextually defined modal bases, ordering sources, existential vs. universal quantification.

The Hacquard move is important because it moves the formal apparatus into a position where it can operate in a way that makes it sensitive to syntactic context. It does so by making the contextual variable that determines the modal base bindable from within the sentence, depending on height of attachment of the modal. Crucially to do so, she has to first modify the system to make it that it is an ‘event’ which provides the base, not a world.

For the system being built up and defended in this monograph, the semantic type of a constituent varies systematically according to the particular zone we are operating in. These zones also correspond to the three main types of modal

meaning. It seems natural therefore to try to derive the differences in interpretation directly from the *semantic properties of the complement* of the modal, rather than via the binidng of a contextual variable as in Hacquard's system. This is the task I turn to next, for the circumstantial modals.

5.5 Circumstantial Modals as Modifiers of Spatiotemporal Properties

In the system being proposed in this book, spatiotemporal properties of the event being described are the elements denoted in the second, inflectional domain of the clause. I assume that a spatio-temporal relation between an event and an reference/perspectival situation also includes a specification of possible world. Thus 'accessibility' relations, coded as R in the Portner schema, are in effect subtypes of f (properties of the event - anchor relationship), properties of which are the business of the second phase. It will make sense then that circumstantial modality in general is expressed in this domain. In contrast, dynamic modality seems to refer to inherent or intrinsic properties of the agent with respect to a particular event, and it will make sense to try to build these meanings out of the first phase. Finally, epistemic modality is about the knowledge and choices of the speaker in seeking a truthful assertion. We should try to make that fall out from the fact that they are attached even higher in the functional sequence than circumstantial modality. Once again, ideally, we would like a unified denotation for modal meanings that underpins all three types of event modulation.

To preserve the insights of the classical model within the present system, we need to find a way of representing the flexibility of the modal base and tying it to height of attachment. We have assumed so far that the inflectional domain, the second phase, is characterized by the fact that it builds properties of anchored event properties. I repeat the denotation for something of the AspP type in (54).

$$(54) \quad [[\text{AspP}]] = \lambda f \lambda d \exists e [\text{Utterance}(d) \wedge f(d, e) \wedge \ldots \wedge u \ldots (e)]$$

In the case of the perfect, the merge of an auxiliary introduced a new situational variable which constituted an intermediate reference situation for the ultimate anchoring relationship for the first phase event. Specifically, *have* introduced a reference situation s' which was related to e and which was then an intermediary in the anchoring of e to the utterance situation d .

$$(55) \quad [[\text{have}]] = \lambda Q \lambda x \lambda f' \exists s' \lambda d \exists f [Q(f)(d) \wedge \wedge \text{State}(s') \wedge \text{HOLDER}(s') = x \wedge f = \lambda s \lambda d [s' \text{ gives evidence for a spatio-temporal relation between } s \text{ and } d \text{ in the same world as } s'] \wedge f'(s')(d)]$$

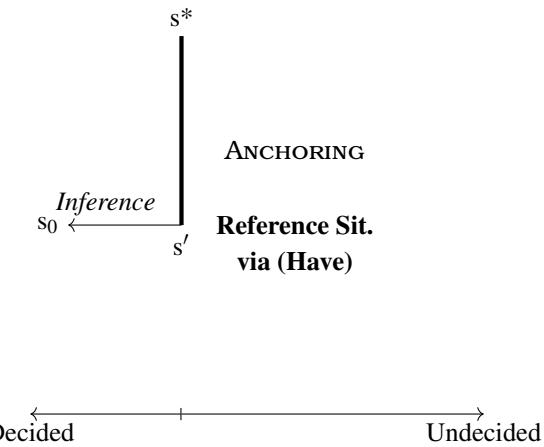
For the circumstantial modal auxiliary, we will assume something similar. The modal will introduce a perspectival situational variable s' with respect to which the situation denoted by the first phase s_0 is oriented. While the perfect expresses an inferential relation between the reference situation and the prejacent situation, the modal will express a projective, predictive relation between the reference situation and the prejacent.

The parallelism between the modal auxiliary and the perfect auxiliary is illustrated in the following figures.

In figure (8), we see the schematic relationship for the perfect. The perfect auxiliary introduces an s' reference situation which is the outer situational variable which will be eventually anchored to tense. The relationship between s' and s_0 is evidential/inferential and s_0 is in the determined past with respect to s' .

Figure 5.11

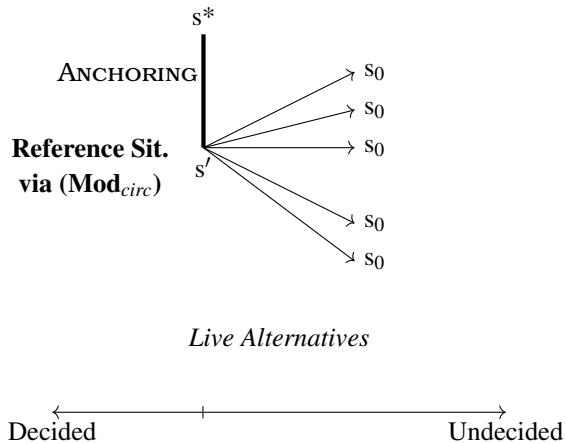
Schema for The Perfect



In figure 5.11, we see the corresponding picture for the circumstantial modal. The modal auxiliary also introduces an intermediate, or reference situation (or perspectival situation) which is the outer situational variable eventually subject to anchoring. But now the relationship between s' and s_0 is projective instead of inferential.

Figure 5.12

Schema for Circumstantial Modality



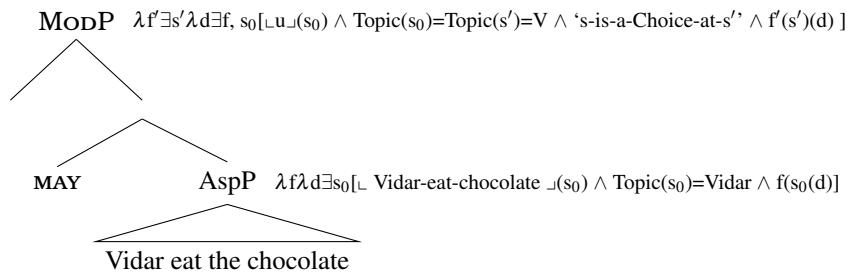
Comparing the two figures, we can see that it expresses a metaphysical modal base schema (as described in Werner (2006) and Condoravdi (2002)) in which worlds up to a given time are strictly identical (the actual world) and only diverge in the future of that given time. Within the metaphysical scheme, the perfect asserts the base situation to be in the determined past of the reference situation; the circumstantial modal asserts the base situation to be in the projective nondetermined future of the reference situation.

For circumstantial modality, therefore, I will be expressing the primitive notion of circumstantial possibility in terms of ‘live alternatives’. Circumstantial live alternatives are simply different ways in which the future could turn out. A circumstantial live alternative, is a situation with particular time and world parameters that is still ‘up for grabs’. It cannot therefore have a time specification that is either present or past with respect to the perspectival situation. The modal combines with the constituent expressing properties of spatiotemporal properties of s_0 anchored at d , and states that such an f exists as a vector expressing a live alternative from s' . The denotation for the circumstantial modal will therefore look as in (56) below.

$$\begin{aligned}
 (56) \quad [[\text{Mod}_{\text{circ}-\text{may}}]] &= \lambda Q \lambda x \lambda f' \exists s' \lambda d \exists f [Q(f)(d) \wedge \wedge \text{State}(s') \wedge \text{HOLDER}(s')] \\
 &= x \wedge f = \lambda s \lambda d [s \text{ is located at a world-time pair that is a CHOICE for} \\
 &\quad \text{the perspectival topic in } s'] \wedge f'(s')(d)
 \end{aligned}$$

In words, this says that there is a perspectival stative situation s' for which s_0 (the one characterized by the event property described in the first phase) is

a live alternative for the topic argument. It is this perspectival stative situation that will eventually be anchored by the tense predicate.



In the tree, we see the sentence ‘Vidar may eat the chocolate’ where *may* is merged in the second phase. The meaning built up is that ‘eating the chocolate’ is a live alternative for ‘Vidar’ in the present situation. The specific lexical item *may* contributes the presuppositional content that the reason that this is a possible alternative for him is because of the permissions that have been accorded him.

The spatio-temporal relationship between the perspectival situation and the prejacent situation performs intuitively the same role as the Kratzerian accessibility relation in terms of possible worlds and it represents a vector distance with respect to a time and world from s' which is a ‘possiblity’ for the perspectival topic in s' .

Looking at it under this configuration, it seems that the relationship that unifies the usage of various different modals is the abstract notion of CHOICE among live alternatives. Intuitively, Topic x is in a situation s' in which she is faced with a set of live options or potentials. The modal expresses the idea that there is some degree of uncertainty in the characterization of the situation (as is the case in all instances of future prediction) and it specifies the prejacent as a CHOICE available to x in s' .

Box 5.2
Informal Schema for Modal Denotations

A modal meaning involves the assertion of a CHOICE within a set of ‘live alternatives’ for a Topic individual x in a perspectival situation s' . These alternatives are directly constructed from the constituent that the modal attaches to.

If AspP has built up a property of event properties based on the existence of an *e* of a particular type, the effect of the CHOICE modal is to assert that the pivot argument is free to choose a situation that instantiates the event *e*. The reasons for that freedom can be permissions, or circumstantial possibilities, internal abilities. This I assume is contributed by the individual lexical presuppositions of the particular circumstantial modal. In this system, the circumstantial alternatives generated are not affected by the modal's own presuppositions, but are the same regardless of the circumstantial modal in play. The fact that they are constructed at the level of the second zone means that 'live alternatives' means that we are talking about values of *f* that are 'still in play' or have not been decided. The fact of generating live alternatives at the level of properties of *f* is what constrains the alternatives to be in the future of the perspective situation introduced by the modal.

The proposal here is intended to seriously build in the idea that modal assertions are made from a background of uncertainty, where there are a number of different potential ways the world could be, but where a particular choice among those potentials is being asserted as a possibility.

Moreover, as discussed in section 5.3, because of the polarity sensitivity of modals, we also need to assume that the modal sometimes combines with PolP in the syntax and that intuitively, both positive and negative alternatives are also explicitly members of the set of 'alternatives' that are live from the perspective of *s'*.

Solving all the problems of representing negation in a model like this is beyond the scope of this monograph. I will simply stipulate here what seems to work in the context of the empirical facts. Recall that I have assumed that clausal (polarity) negation lies within the second phase and therefore combines with properties of spatiotemporal properties. I will assume therefore that polarity negation simply negates the existential quantifier binding the event variable that has been closed at Asp.

$$(57) \quad \begin{aligned} [[\text{PolP}_{\text{pos}}]] &= \lambda f \lambda d \exists e [\text{Utterance}(d) \wedge f(e)(d) \wedge \sqsubset u \sqcup(e)] \\ [[\text{PolP}_{\text{neg}}]] &= \lambda f \lambda d \neg \exists e [\text{Utterance}(d) \wedge f(e)(d) \wedge \sqsubset u \sqcup(e)] \end{aligned}$$

When it comes to the assertion of a simple CHOICE for the pivot argument, whether the syntactic category of the complement is AspP or a PolP is immaterial to the circumstantial alternatives generated. The circumstantial alternatives will be all the logically possible alternative states of affairs fanning out from the perspectival situation at the relevant times in the future. These will inevitably include some situations in which the prejacent eventuality takes place and some in which it doesn't.

The only difference will lie in whether polarity negation can scope outside of the modal meaning and negate it directly or not. If a modal selects AspP rather than PolP, then polarity negation will be interpreted as negating the modal. This means that if one wanted to assert the circumstantial possibility of *not* doing something, then low constituent negation, and not polarity negation would have to be used. I assume that negation, like modality itself has three zonal manifestations. I give an informal characterization of what I have in mind here in (58).

(58) (a) **Constituent negation**:

for verbs is within EvtP, the event essences domain and constructs an actual eventuality which manifests positive properties that are the negation of the description deployed, an anti-event, if you will.

(b) **Polarity negation**

is in the spatiotemporal properties domain and introduces a negative operator on the existential binding of e

(c) **High negation**

is in the speech act zone and gives rise to a negation of the accuracy of particular deployed lexical item (subject to focus).

In fact, there seems to be very little difference when it comes to deontic modals between the assertion of the permissability that an event does not exist in the future, and the assertion of the permissability of an event that positively displays the negated properties described. This is because the anti-event is always stronger in this kind of permission context. If instantiating the anti-event is allowed then it entails that merely not instantiating the event with those descriptive properties is also allowed. This is because the two scenarios, to the extent that they differ seem to belong to the following kind of implicational scale of decreasing likelihood of being permitted.

(59) Mary doesn't go to the party > Mary makes a big dramatic point of visibly failing to show up at the party.

If you are allowed something on a particular point of the scale, then you are allowed everything higher than it as well.

So the difference between AspP selecting deontics and PolP selecting deontics does not give rise to a serious difference in expressivity when it comes to this aspect of the meaning of the modal. On the other hand, if a modal does select PolP, then polarity negation can only be interpreted as negating the event variable. In that case, negating the possibility itself would have to be handled by periphrasis. This seems like an expressive disadvantage. It is for this rea-

son, I think, that deontic ‘existential’ modals tend to scope under negation, i.e. they tend to be non PolP selectors. Thus, purely circumstantial ‘existential’ modals do not tend to be PolP selectors in natural language quite generally. As Iatridou and Zeilstra (2013) point out, languages do not seem to possess existential deontic modals that embed negation. The availability of low negation for the one reading, and the clear communicative pay offs from being able to negate the permission itself via Pol is probably what leads to this pattern.

The other point to consider here is that existential circumstantial modals often derive from a dynamic version which lives inside the Evt domain, in the case of *can* this is the ability reading. I will assume that the dynamic verb is a species of *init*, and like lexical verbs never scopes over any kind of negation at all—it is merged too low to scope over even low zone negation. Empirically, it seems that if a modal has more than one zonal use, then the selection of a negation/position projection or not carries over to *all* of its uses. Basically, this means that if there is a dynamic (low) use for the modal at all, then the scope with respect to negation will be consistently fixed as being *under* negation.

Before we give denotations for the different circumstantial modals in English, we still need to capture the difference between ‘possibility’ and ‘necessity’ modals in this kind of implementation. The existential binding of the *f* property relating *s'* to the prejacent as a valid live alternative is tantamount to existential assertion of possibility. The innovation we need here is not to introduce universal quantification over *f* for so-called universal modals, but think of the circumstantial necessity modals *non*-quantificationally in terms of *exclusive* choice. In the formula below, I have replace the simple CHOICE predicate with the one that goes with necessity modals—the unique EXCLUSIVE CHOICE.

(60) (a) Jane must_{[PolP} [AspP < Jane > sing]].
 (b) $[[\text{Mod}_{\text{circ-must}}]] = \lambda Q \lambda x \lambda f' \lambda s' \lambda d \exists f [Q(f)(d) \wedge \wedge \text{State}(s') \wedge \text{HOLDER}(s') = x \wedge f = \lambda s \lambda d [s \text{ is located at a world-time pair that is the ONLY CHOICE for the perspectival topic in } s' \wedge f'(s')(d)]$

Here once again, the Set of Alternatives is: all the different possible values of world time vectors projected from *s'*.

When it comes to the ‘necessity’ circumstantial, which express exclusive CHOICE, the difference between being able to select a negative PolP vs. a positive PolP gives quite distinct directive effects. Moreover, compelling a pivot to perform an anti-event does not entail that you would be content with them merely failing to perform an event of that type. In the scale set up in (59), if you are forced to do the higher thing, then doing the lower thing is also fine, essentially reversed from the simple CHOICE case. Thus, low negation does not

give the same effects for the exclusive CHOICE deontic modals. ‘Necessity’ circumstantial can therefore in principle come in either selectional setting.

In the following table, I give a listing of the circumstantial modals of English and their selectional setting. Mod^{Pol} indicates that the modal in question selects for PolP and therefore always scopes over polarity negation when it exists; ^{Pol}Mod indicates that the circumstantial modal in question selects for a non polarity marked AspP and will therefore always scope under polarity negation when it exists. Following Copley (2002), I will assume that the modal *will* in English is a modal in the circumstantial zone (as opposed to being a tense, or even an epistemic modal). The reason I think that future *will* is not an epistemic modal, but rather a circumstantial modal of prediction is that it is the purest instantiation of circumstantial live alternative projected into the future. In fact, there *is* an epistemic version of *will* which speakers spontaneously produced and it is crucially different from future *will* in that it is anchored to the speech time and reports on the speakers predictions about current state of affairs based on their knowledge (61).

(61) (a) John will go to the party. *predictive circumstantial modal*
 (b) John will be in his office now (because I know his habits). *epistemic*

Including *will* therefore, here is the table for circumstantial modals in English, notated for polarity selection, for presuppositional flavour, and for zonal variants.

| Circ. Modal | \pm Exclusive CHOICE | Presuppositional Flavour | Zones |
|-----------------|---------------------------|-----------------------------|-------------------|
| ^{Pol}Can | — | facilitative | (ep) < circ < dyn |
| ^{Pol}May | — | permission | circ |
| Will Pol | + | prediction | ep < circ |
| Should Pol | + | normative | ep < circ |
| Must Pol | + | directive | ep < circ |
| $^{Pol}Could$ | — | facilitative | (ep) < circ < dyn |
| $^{-Pol}Need$ | + | non-obligation | ep < circ |

Finally, we observe that modal verbs in English combine with the bare ‘uninflected’ form of the verb. We have already seen that the bare uninflected root in English can lexicalize the full complement of verbal heads, and thus,

by hypothesis lacks T, or Asp information. Furthermore, it is the form that is morphologically eligible for suffixation. I will assume that the form the combines with modal verbs is this very same item, suffixed with an irrealis Asp head, which represents the infinitival ending, which is null in English.⁵⁸

The system I have outlined proposes analysing modals as expressing CHOICE over linguistically constructed LIVE ALTERNATIVES. But what is a LIVE ALTERNATIVE? Arguably, it is nothing more than ‘possibility’ itself relativized to a particular situational reality and topic. It has the same status as the notion of necessity or possibility in the axiomatization of modal logic, or in the status of the notion of what is ‘possible’ in the construction of possible worlds. It is the irreducible axiomatic part of the idea of potentials or hypotheticality in the domain of language. We have already established that human users of language employ a dimension of this kind in their daily expressivity, and it needs to be represented as one of the compositional building blocks. The only difference between the above system and the classical treatments lies in *where* the primitive is located in the axiomatization of the logic that modal meanings builds on. In this particular treatment, the quantificational analogy is rejected and the system is built from an analogy to CHOICE, generalizing from the deontic core cases instead. The primitive CHOICE relation asserts the freedom of the ‘pivot’ within the space of hypothetical alternatives, from a particular situational vantage point. Simple CHOICE does nothing more than assert a particular freedom for the pivot, given the GROUNDS for constraining ones options contributed by the presuppositional content of the modal.

The ‘pivot’ for the CHOICE is a crucial argument of the CHOICE relation, and is the topic (either explicitly or implicitly) of the perspectival situation. Something can be a choice for a pivot x if it is part of the things x is able to do, is allowed to do, or is logically possible for x to do. The important thing is that the choices are relativized to the involvement of x. In circumstantial modals, this x argument is usually the highest or external argument of the event in the situational description, but it can also be other arguments or even filled in contextually.

There is an important role in these definitions for different presupposed information or pragmatic contextual information about the GROUNDS for why the pivot has the CHOICE he/she/it does. This framework is not intended to replace the contextual input to modal semantics. the ‘Grounds’ for a CHOICE are in part contributed by the lexical presupposition of the modal itself and in part by linguistic context and other contextual factors.

⁵⁸ The information of this null morpheme is to provide the presupposition that the verbal eventuality in question is located in a different world than the world of the perspective situation.

(62) Grounds for CHOICE coming from discourse context:
A: Oh no, I have a meeting at 9 a.m. tomorrow morning!
B. Then, you must get up before 8 for once.

(63) Grounds for CHOICE coming from adverbial modification:
If you want to make that meeting, you must get up before 8.

5.6 Conclusion

In this chapter, the specific meaning contribution of circumstantial modals turned out to be similar in an interesting way to the function of the *have* auxiliary in English, which also resides in this domain. As discussed in the previous chapter, the *have* auxiliary was argued to introduce a proxy or intermediate situational variable s' bearing a particular relation to the core situational description s_0 . This intermediate situational variable was equivalent to the one that has been called the reference situation in linguistic work on aspect (Reichenbach 1947; Klein 1994 *inter alia*). In the view of this chapter, it is also equivalent to the perspectival situation invoked in work on modality (Condoravdi 2002). The intermediate situational variable is the one that is eventually anchored to tense. It is a special case of the spatiotemporal property of s_0 , the core situational variable existentially bound at the edge of the first phase, and which is then anchored only indirectly to d . Like perfect auxiliaries, circumstantial modals introduce an intermediate situational variable in relation to the core event. However, instead of being in a relationship of entailment to s_0 , s' is in a relation of prediction. Specifically modals assert choice among live alternatives as part of their lexical semantics. They divide into the ones where the choice is simple CHOICE among many (the so-called ‘existential’ modals), and the ones where the CHOICE is assessed to be the *only* option the is faced with (the ‘universal’ modals).

In changing the compositional semantics for modals away from the classical model, the idea is not to deny the different components of meaning isolated in that tradition or to deny the work done by contextual factors. Rather, in changing the architecture, we get the kind of mapping between the syntax and semantics that will allow us to integrate modality more naturally with temporal interpretation, and to unify the modals’ denotations across domains. In the following table, I compare the classical system with its parameters of (i) quan-

tificational force (ii) modal base and (iii) ordering source (as in Kratzer 1977) with the proposal made in this chapter.

| | CLASSICAL MODEL (Kratzer 1977) | CHOICE Semantics for Modality in QQS |
|--|--|---|
| Quantificational force | \exists | CHOICE |
| | \forall | EXCLUSIVE CHOICE |
| Modal Base (Primary effects) | Dynamic Circumstantial Epistemic | Syntactic Height within EvtP Syntactic Height between EvtP and TP Syntactic Height above TP |
| Modal Base (Secondary) plus Ordering Source | deontic, buletic teleological etc. | presuppositional content from modal concerning nature of GROUNDS for CHOICE |
| Modal Base (Secondary) | contextual and linguistic factors constraining modal base | contextual and linguistic factors on GROUNDS for CHOICE |

In the next chapter, I will extend the account to epistemic interpretations which by hypothesis are built higher up in the phrase structure. In addition, we need to establish what happens at the level of T and how situational anchoring is achieved in this system.

6 Modals and Generalized Anchoring

In the previous chapter, we introduced the issue of modal meanings and provided an account of how circumstantial modal interpretations are captured in a framework which models the second phase of the clause as denoting spatio-temporal properties of event properties. We saw that like the auxiliary *have*, circumstantial modal meanings are distinguished by the fact that they introduce an intermediate reference situation, which in the case of modals, is the *perspective* with respect to which the most deeply embedded event is located. While the auxiliary *have* locates the embedded event in its own ‘realistic’ past, the modal auxiliaries introduce a reference situation that has the embedded event in its own ‘predicted’ (i.e. non-real) future. In this chapter, we focus first on another aspect of modal meaning that is an important semantic component of the extended verbal projection, namely temporal anchoring.

Whatever situational description has been built up in the second phase, whether simple (no intermediate reference situation), or auxiliated (with intermediate ‘reference’ situation), the outermost situational variable needs to be explicitly related to the contextual anchor point, the speech time, in order to create something that has actual truth conditions. This is classically seen as the job of tense, but in modal constructions, it is the modal itself that occupies the T position and therefore by hypothesis, also contributes information related to anchoring. In what follows, I summarize the patterns of anchoring interpretation found with modals in English. It should be clear from the discussion that follows that anchoring information is indeed modal-specific in an important way. Once the patterns are established, I will propose an account of the contribution of modal anchoring in terms that mirror the anchoring contribution of tense, and make a proposal for how that meaning is integrated into the denotations we have been building up so far.

Once this is complete, I will turn finally to epistemic modal interpretations, and make a proposal for how these are constructed from the basic building blocks we have already seen. To anticipate the claims of this final section, I

will argue that epistemic meanings emerge precisely when they attach to constituents whose temporal parameter has already been fixed. In addition, they will not introduce a new reference situation, but use the utterance situational variable d as their ‘perspective’ situation.

6.1 Temporal Properties of Modals

There are interesting generalizations that emerge from a careful description of the relationship of the modal assertion to temporal properties, and it is this that I turn to now, taking my lead from Condoravdi (2002).

Condoravdi (2002) points out clearly the need to distinguish between the temporal perspective of a modal and its temporal orientation. In the system we are building up here, the perspectival situation is our s' , introduced by the modal itself. The event time is the spatiotemporal position of the embedded event e , as built up by the first phase. If we consider the following three examples of modal sentences, we can see that in (1-a) the event of ‘going to the party’ is in the future from the time of utterance, in (1-b), John ‘being in his office’ is contemporaneous with the time of utterance, and in (1-c) the event of ‘winning the race’ is previous to the time of utterance.

- (1) (a) John might go to the party.
- (b) John might be in his office.
- (c) John might have won the race.

One way to capture this is to specify the semantics for **MIGHT** directly in two versions, one forward shifting and the other non-shifting with respect to the possible worlds considered. One might even stipulate a third composite modal **MIGHT-HAVE** which requires the possible worlds to precede the utterance time. Condoravdi (2002) gives the following three possible denotations corresponding to such a view (although, as will become clear, she herself proposes a more compositional treatment and unified conception of *might*). In the following **MB** designates the modal base that a modal depends on for its interpretation (Kratzer 1977).

- (2) (a) Forward shifting modals:
 $\text{MIGHT}_{MB}^1 \phi$ is true at $\langle w, t \rangle$ iff there exist w', t' such that $w' \text{ in } MB(w, t), t \prec t'$ and ϕ is true at $\langle w', t' \rangle$. (b) Non-shifting modals:
 $\text{MIGHT}_{MB}^2 \phi$ is true at $\langle w, t \rangle$ iff there is w' in $MB(w, t)$, such that ϕ is true at $\langle w', t \rangle$. (c) Backward-shifting modals:
 $\text{MIGHT-HAVE}_{MB}^1 \phi$ is true at $\langle w, t \rangle$ iff there exist w', t' such that $w' \text{ in } MB(w, t), t' \prec t$ and ϕ is true at $\langle w', t' \rangle$.

These three interpretations all take the present utterance time as the perspective, but this too is independently modifiable. In certain contexts the perspective of the modal can be shifted backwards, to say that at a particular point in time in the past, a modal statement was true. These cases are particularly clear when we look at Dynamic modality, where intuitively the moribund past tense on *could* actually seems to do some transparent semantic work (3).

(3) In those days, John could easily swim 2 kilometers.

But past perspective for the modal is also possible in embedded sentences for (future-meaning) *would* (4-a) and (epistemic) *might* (4-b), and the circumstantial version of *could* (4-c).⁵⁹

(4) (a) Last year, John told me that he would quit his job.
(b) Last year, John told me that he might quit his job.
(c) Last year, John told me that he could take vacation any time he wanted.

To distinguish between these two different aspects of the temporality of modal meaning, Condoravdi uses the terms *Perspectival Time* (PT) and *Evaluation Time* (ET). The Perspectival Time is the time at which the *potential* for the prejacent event is asserted, and the Evaluation time is the time of the prejacent event itself. This corresponds directly, as I have said, to the perspective situation s' and the embedded situation s_0 respectively. We can see already that it is not a simple matter of inspecting the morphology or the lexical items in question to know how PT and ET are configured in any particular case.

Condoravdi (2002) provides some important generalizations with respect to this patterning, which I demonstrate here. Firstly, deontic modality always forward-shifts, even with stative predicates. Consider the examples in (5) using *can* with the meaning corresponding to circumstantial possibility.

(5) (a) John can go to the party (if he does his homework). *Forward-shifted*
(b) John can be in London by noon (if he takes the early flight). *Forward-shifted*

⁵⁹ See Stowell 2004 for a discussion of these facts and an argument that moribund past tense morphology on English modals actually is grammatically interpretable.

Epistemic modality is variable, but is sensitive to *akitionsart*: dynamic predicates induce forward shifting while stative predicates produce non-shifting readings.⁶⁰

(6) (a) John might go to the party, but I wouldn't count on it. *Forward-shifted*
 (b) John might be in his office, but I wouldn't count on it. *Non-shifted or Forward-shifted*

In my proposal for circumstantial modality in the previous section, the forward shifting property of circumstantial modals arises obligatorily because the construction of the notion of **LIVE-ALTERNATIVES** which simply are not defined for times simultaneous with or previous to the perspectival situation.

Finally, back-shifting seems to be possible only in the context of additional linguistic material: the addition of the perfect auxiliary *have*, or the embedding under a past tense matrix predicate. In the former case, we find only epistemic modals.

(7) (a) John must have won the race.
 (b) John might have won the race.
 (c) John could have won the race.
 (d) John may have won the race.

In the latter case, we find deontics and dynamics in addition to epistemics, but only a lexical subset of them (basically those with moribund past tense morphology).

(8) (a) John said that he would go to the party.
 (b) John said that he could swim 2 kilometers.
 (c) John said that he could take vacation whenever he wanted.
 (d) John said that he might go to the party.
 (e) John said that I should go to the party.
 (f) ?John said that he must go to the party.
 (g) ?John said that he may go to the party.

⁶⁰ What Condoravdi actually says is “The correct generalization is that modals for the present have a future orientation optionally with stative predicates and obligatorily with eventive predicates.” She claims further that this fact is independent of the flavour of modality in question. I have reason to doubt this latter claim and relativize her statement to epistemic flavours. In the case of circumstantial modality, it seems to me that stative predicates obligatorily forward shift just like dynamic ones.

- (h) ?John said that he will go to the party.
- (i) ?John said that he can go to the party.

In the final four sentences of (8) have been given question marks— they seem to be good only on the equivalent of a kind of ‘double access’ reading for the modal perspective.

The generalizations that we have from examining the interaction of English modal meanings with temporal information and *akionsart* can be summarized in the following two figures. Given the agenda of this monograph, these patterns should be made to fall out from a compositional treatment of modal meaning combined with the properties of the constituents they combine with.

I separate the generalizations concerning perspectival time from those concerning evaluation time. In the case of evaluation time, the generalizations depend on the type of modality and the *akionsart* of the prejacent (see Figure 6.1).

Figure 6.1
Evaluation Time

| | <i>Forward-shifted</i> | <i>Non-shifted</i> | <i>Backward-shifted</i> |
|-----------------------|------------------------|----------------------|--------------------------------|
| <i>Epistemic</i> | YES | YES (States only) | YES (With <i>have</i> only) |
| <i>Circumstantial</i> | YES | NO | NO |

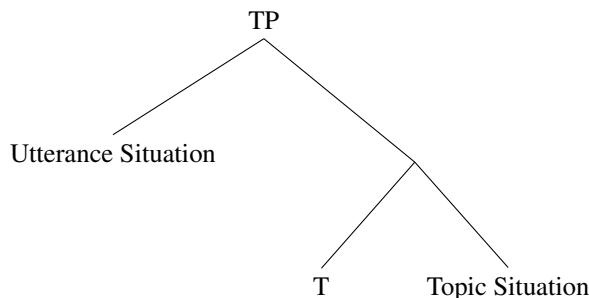
In the case of perspectival time, the generalizations depend on the particular modal and the kind of morphology it possesses (see Figure 6.2).

Figure 6.2
Perspective Time

| | <i>Present</i> | <i>Past</i> |
|---------------|----------------|-----------------------|
| <i>must</i> | YES | NO |
| <i>may</i> | YES | NO |
| <i>can</i> | YES | NO |
| <i>will</i> | YES | NO |
| <i>might</i> | YES | YES (under embedding) |
| <i>could</i> | YES | YES (under embedding) |
| <i>should</i> | YES | YES (under embedding) |
| <i>would</i> | YES | YES (under embedding) |

Under the classical view of anchoring, the T head combines with a situational description (the Topic situation) and establishes a relationship between it and the utterance situation (see Figure 2). I will call this anchoring, and assume that tense relationships are just one possible instantiation of the anchoring relation (see Ritter and Wiltschko 2009).

Figure 6.3
TP as the locus of generalized situational anchoring



In the present framework will be designed to express the very same basic intuition. The topic situation is the situational description provided by the main verb (in the case of non-auxiliated structures) or by the reference situation introduced by an auxiliary. I will then assume, following the intuition of the classical model, that the time parameter of the topic situation is constrained by the establishment of a relationship between it and the time parameter of *d*. This relational information is what is assumed to reside in *T*, and I will continue to use the label *T* for the node that contributes this relation. This is thus a straightforward extension of neo-Reichenbachian theories of tense/aspect whereby a reference time or topic time (see Klein 1994 and Demirdache and Uribe-Etxebarria 2008) is the intermediary between the utterance time and the event, with the difference that I only invoke intermediate situations in the case of auxiliary structures.

In what follows, I will also use the * notation to indicate the parameters of the utterance situation *d*, t^* = anchor time and w^* = anchor world.⁶¹

As is uncontroversial, the temporal predicate *PAST* is a one that establishes a temporal relation of precedence between the topic situation (the situational description denoted by the complement to *T*) and the utterance situation.

⁶¹ This at least is the assumption for normal matrix situations, but I assume that this can be relativized to deal with embedded attitudes and free indirect discourse.

I will assume that **ANCHORING** is simply the existential closure of spatiotemporal location for the outermost situational variable, achieved by relating it explicitly to the utterance situation. The proposition is thus identical to the existential closure of the f_d variable. What results is a constituent that denotes a property of current utterance. The utterance d , has arguments Speaker and Hearer that, like the utterance situation itself are indexically bound.

Constituents larger than TP in the clause will be properties of the utterance situation and will be written according to the following schema, for a particular situational description s by the second phase:

(9) $\lambda d \exists f \exists Q [Q(f)(d) \wedge \text{Source}(d) = \text{'Speaker'} \wedge \text{Goal}(d) = \text{'Hearer'}]$
 Where Q stands for the predicate of situational properties already built up by the second phase.

To give a concrete example without auxiliaries, consider the simple sentence *Vidar ate the chocolate*. We assume that the final AspP (maximal constituent in the second phase) for that sentence has the denotation:

(10) $[[\text{AspP}]] = \lambda f \lambda d \exists e [\text{Utterance}(d) \wedge f(d)(e) \wedge \perp \text{Vidar eat chocolate} \lrcorner (e)]$

The temporal predicate **PAST** contributed by the morphology here, anchors the situation to the utterance as shown in (11).

(11) $[[\text{TP}_{\text{past}}]] = \lambda d \exists f \exists e [\text{Utterance}(d) \wedge f(d)(e) \wedge \text{PAST}(f) \wedge \perp \text{Vidar eat chocolate} \lrcorner (e) \wedge \text{Source}(d) = \text{'Speaker'} \wedge \text{Goal}(d) = \text{'Hearer'}]$

In this framework, the **PAST** predicate must be a predicate over spatiotemporal properties rooted in d , f_d , and we can specify it informally as in (12).

(12) $\forall \text{eventualities } e \text{ and speech events } d, \text{ such that } f(d)(e), \text{PAST}(f) \text{ is true iff the temporal parameter of } e \text{ precedes the temporal parameter of } d.$

In the case of the present tense, I will assume for reasons that will be obvious as we proceed that the present tense contributes the information that the outermost situational variable is anchored to the utterance time via identity, and moreover that the utterance time is abstractly represented as a moment, not as an interval. This is because the English present has the peculiar property that it only combines felicitously with states, and I will continue following Taylor (1977) in assuming that the crucial distinguishing property of states is that they are able to be true at a single moment. Thus for the present tense sentence *Vidar likes sushi*, we would have the denotation in (13)

(13) Vidar likes sushi.

$$[[\text{TP}_{\text{pres}}]] = \lambda d \exists f \exists e [\text{Utterance}(d) \wedge f(d)(e) \wedge \text{PRESENT}(f) \wedge \ldots \text{Vidar like sushi } \ldots (e) \wedge \text{Source}(d) = \text{'Speaker'} \wedge \text{Goal}(d) = \text{'Hearer'}]$$

Where the PRESENT is a property of f, a relation between e and d, defined informally as in (14)

(14) \forall eventualities e and speech events d, such that $f(d)(e)$, PRESENT(f) is true iff the temporal parameter of e is *identified with* the temporal parameter of d, the moment of speech.

In the case of our perfect and modalized sentences, the same T semantics applies—the innermost situational variable only gets anchored *via* the reference situation s' , which is the one that is directly affected by the tense predicate.

Temporal information such as PAST, or PRES in English perform the shift from properties of situational properties, to properties of the utterance. But as we also know, modals in English all behave distributionally as if they end up high in the clause: they invert in questions, they precede negation and do not require *do*-support, and they are unique in the clause. I take these facts at face value and assume that, wherever the modal is actually base generated, it always ends up in the equivalent of Infl. What we assume from tense is that it is the job of this head to establish a relationship between the reference situation s' and the situational anchor of the clause d. The natural assumption then is that modals also must be endowed with information that establishes such a relationship. I assume moreover that this anchoring property of Infl, suitably generalized, is a universal property of natural language sentences, plausibly driven by constraints at the interpretational interface (Ritter and Wiltschko 2009). I will assume therefore that modals also anchor the situation, but do so by relating the situational description to the anchor *world* or time, or both.

To accommodate the modals we will need to generalize our approach to tense slightly. I will assume that anchoring comes in two main flavours, the kind of anchoring we saw with the past tense above where a constraint is placed on the denotation of the temporal interval in relation to the utterance time, and the kind in which the situational is directly indexically *identified* with the utterance situation. In the former case, some kind of discourse or anaphoric binding of the temporal variable is necessary to provide the actual value of the temporal variable. The two types of anchoring relations are defined informally below.

Box 6.1
Types of Anchoring

\forall eventualities e and speech events d , such that $f(d)(e)$, INDEX(f) is true iff the temporal parameter of e is strictly identified with the temporal parameter of d , the speech time.
 \forall eventualities e and speech events d , such that $f(d)(e)$, ANAPH(f) is true iff the temporal parameter of e is resolved *anaphorically*, either by binding from something in the linguistic context, or to some purely discourse contextual topic time or world.

Under this view, the past tense in English is in the ‘anaphoric’ camp, while the present tense is clearly indexical.⁶² The idea that modals in some sense do the same job as tense has echoes in the proposal found in Iatridou 2000 (see also Isard 1974) in which past tense morphology is not a primitive tense category, but is one of the manifestations of the more general semantic category (REMOTE, in their terms). According to Iatridou (2000), the remote relation can relate worlds as well as times, accounting for some cases of past morphology on modals. My proposal is different from Iatridou’s in that I take INDEXICAL vs. ANAPHORIC to be the primary relational distinction, not IDENTITY vs. REMOTE, although I think the guiding intuition is the same.⁶³

There is also a tense auxiliary in English, *do* and a monograph on English auxiliaries would not be complete without some discussion of it. I assume that *do* can be inserted directly in T in English to spell out the temporal features anaphoric and indexical (*did* and *does*) respectively. However, the conditions under which this is allowed and/or forced in competition with the simple past tense would take us too far afield here.

Turning then to anchoring by modals, I will argue that they too come in two distinct varieties depending on the relationship they establish between the perspectival situation s' and the anchor context d : the indexical and the anaphoric.

Recall from the discussion on the tense interpretation of modals in Condoravdi (2002), we noted two strong generalizations. One was that circumstantial modality was obligatorily forward-shifted while epistemic modality did not

⁶² I put aside other interpretations of the present tense in English here. Under certain discourse conditions, additional meanings for the present include narrative past and planned future. I assume that habitual present involves the construction of a derived dispositional state in the first phase. My suspicion is that narrative past and planned future both involve the implicit building of a reference situation which is truly present, with a forward or backward inferential relation. However, since the correct analysis of these constructions would involve much more data study, I leave it for further work.

⁶³ The difference between the REMOTE and the ANAPHORIC specification arises in cases where for example ANAPHORIC type anchoring results in reference resolution overlapping the speech parameters. This would be disallowed by a REMOTE specification, but fine for a modal specified as ANAPHORIC. My choice of privative features reflects the fact that I think the latter type exists.

have to be. The other was that some modals introduce a perspective time that is obligatorily identified with the speech time, while other modals are more flexible. The hallmark of this extra flexibility is the ability to have their perspective situation identified with an attitude in the matrix under embedding, as can be seen in (15).

(15) (a) Vidar thought that he could win the race.
 (b) Vidar thought that he should get a prize.
 (c) Vidar thought that he would win the race.
 (d) Vidar thought that he might get a prize.

The following modals do not allow this kind of perspectival anaphoricity, and give rise to a feeling akin to the double access reading when present tense is embedded under past (16-e) .

(16) (a) ? Vidar thought that he may win a prize.
 (b) ? Vidar thought that he must win a prize.
 (c) ? Vidar thought that he can win the race.
 (d) ? Vidar thought that he will win the race.
 (e) ? Vidar thought that he is deserving of a prize.

The idea here is that modals like *must* share a property with the present tense in being indexically bound to the utterance time. The other modals, in (15), are anaphoric in the sense defined above. Note that anaphoric reference in this sense covers many different modes of reference resolution (being essentially negative); it is intended to mirror the cut in the pronoun system between indexical forms like *I/you* on the one hand, and non-indexical ones like *he/she/it* on the other. It is well known that this is the cut that is almost universally instantiated crosslinguistically within pronoun systems. It means that the actual reference assigned to the temporal variable can be achieved by binding at the discourse level or within the sentence, but crucially reference needs to be resolved since it is not automatically identified with a parameter of the context as it is in the indexical case.

Thus, in addition to its semantics of the existence of ‘live alternatives’, modals in English also possess information that anchors the external situational variable. Some modals like *must* have indexical anchoring specification, while other such as *might* have anaphoric anchoring information. This distinction cuts across the epistemic/circumstantial divide. Thus, we must now augment our modal denotations with this extra information. Since, *must* is an indexical modal, it will provide information anchoring the perspective situation obligatorily to the speech situational parameters. The following would be the deno-

tation of the TP headed by the circumstantial *must* (abstracting away from the argument structure information, for now and for the CHOICE component of the epistemic modal's meaning).

$$(17) \quad [[\text{TP}_{\text{must}}]] = \lambda d \exists f' \exists s' \exists f \exists s_0 [\text{State}(s') \wedge \wedge f(d)(s_0) \wedge \perp u \perp(s_0) \wedge f'(d)(s') \wedge \text{INDEX}(f') \wedge \dots]$$

This means that that the base perspectival situation introduced by *must* is always the same as the utterance time. On the other hand, *might* is not indexical in this sense, but anaphoric, and the temporal and world variables in the perspectival world may be anaphorically resolved to a salient discourse interval, or identified with some linguistically present time via binding.

$$(18) \quad [[\text{TP}_{\text{should}}]] = \lambda d \exists f' \exists s' \exists f \exists s_0 [\text{State}(s') \wedge \wedge f(d)(s_0) \wedge \perp u \perp(s_0) \wedge f'(d)(s') \wedge \text{ANAPH}(f') \wedge \dots]$$

Returning to embedded contexts, as we might expect, anaphoric modals pattern with the simple past tense in English in allowing the time of the embedded situation to be bound by the matrix situation.

(19) John: "I can go to the party."
 John said that he could go to the party.

(20) John: "I may go to the party."
 John said that he might go to the party.

(21) John: "I will go to the party."
 John said that he would go to the party.

On the other hand, the indexical modals *can*, *may*, *and will* themselves, and the singleton *must* are a little strange in embedded contexts and have the flavour of a 'double access' reading, much like the English present tense when embedded under past (22).

(22) (a) John said that he can go to the party. (ability must still be current)
 (b) John said that he may go to the party. (permission must still be current)
 (c) John said that he will go to the party. (party cannot have happened yet)

(d) John said that he must go the party. (obligation must still be current).

A set of questions naturally arises here concerning the behaviour of QQS in embedded contexts, and in particular, the reference of the embedded d eventuality. I will assume that quite generally there are two options for the d variable of an embedded sentence: it can either be identified with the higher d of the matrix verb, giving rise to indexical effects, or by the situational variable of the matrix verb itself if it happens to be a verb of communication. While this mimics in broad outlines the possibilities for anchoring assumed in the literature, an exploration of the specific consequences and predictions of this system as compared to others when it comes to indexical shift will have to be left for further research. It is worth noting here in passing though, that QQS, possibly uniquely, predicts a preferential status for verbs of *saying* in allowing indexical shift (as opposed to attitude predicates more generally).

6.2 Epistemic Modals as Modifiers of Assertions

In the previous chapter, I proposed denotations for the possibility and necessity modals in the realm of circumstantial modality, and I have now been explicit about the anchoring properties of both tense and modals. The outer situational variable introduced by the modal meaning is anchored to the utterance via the modal's temporal setting and asserted as being either identified with the utterance situation d or not.

But what about epistemic modality such as in the sentences shown in (23) below.

(23) (a) Jane might be in Edinburgh.
 (b) Jane must be in Edinburgh.

In this case, the notion of uncertainty or potential seems to lie in a different dimension. As Condoravdi (2002) has already noted, epistemic modality involves quantification over 'worlds' that occur at the same time as the perspectival world. Epistemic modality is not necessarily forward-oriented in the way that circumstantial modality is.

I propose that the source of epistemic readings is precisely that the modal attaches *after* the anchoring of the situation to the utterance, and where f has already been resolved and no circumstantial alternatives are generated. Since we have already had to assume that modals carry both an anchoring specification and a modal meaning in English, let us see what happens if the modal merges first at T.

Taking the case of *must* first, if *must* lexicalizes T it will contribute indexical information with regard to the temporal variable of the situational description constructed up to that point.

According to the denotations given so far, the AspP constructed up to that point, would look schematically as in (24).

$$(24) \quad [[AspP]] = \lambda f \lambda d \exists e [\text{Utterance}(d) \wedge f(d)(e) \wedge \llcorner u \lrcorner(e)]$$

An indexical specification of f here would give rise to the TP in (25)

$$(25) \quad [[\text{TP}_{\text{pres}}]] = \lambda d \exists f \exists e [\text{Utterance}(d) \wedge \text{INDEX}(f) \wedge f(d)(e) \wedge \llcorner u \lrcorner(e) \wedge \text{Source}(d) = \text{'Speaker'} \wedge \text{Goal}(d) = \text{'Hearer'}]$$

Now, we also want the epistemic interpretation of *must* to be related in a systematic way to the interpretation already given for circumstantial *must* in terms of unique CHOICE among LIVE-ALTERNATIVES. Intuitively, the difference between the present tense TP and the epistemic modal meaning actually contributed by *must* is that of a set of live-alternatives not for a topic argument introduced lower down, *but for the speaker herself* because of her incomplete direct knowledge of the facts. This is expressed intuitively in (26).

(26) **Epistemic Must:**

The proposition expressed is the *only* CHOICE for the speaker in the utterance situation d, given the alternatives open to her, consistent with her knowledge.

Crucially, the speaker *is* faced with LIVE-ALTERNATIVES because she has not got complete experiential evidence for what she is about to say, however, what knowledge she does have is consistent only with the assertion made (and not, for instance its opposite). Once described in this way, we see that a defining feature of epistemic meanings is that the ALTERNATIVES with respect to which a CHOICE is being made are not alternatives related to ways in which the world might unfold in the future, but alternatives related to what the fact of the matter is at a particular world and time. The alternatives are at least in part due to ignorance, not to the radical indeterminacy of the future. Also, there is no separate introduction of a perspectival situation here—the perspectival situation is the utterance situation itself, d.

I propose that the reason the epistemic modal can be inserted directly in T and not in the second phase is that one does not need to introduce a reference situational variable to construct the epistemic meaning. The perspectival situation is d itself, and the modal meaning can be expressed as the relationship between d and the embedded situation without intermediary. Or put the

other way round, if the modal does not introduce a novel reference situation by being merged in the spatiotemporal zone, then the CHOICE among LIVE-ALTERNATIVES meaning must be applied to d as the perspective situation.

Since the world and time are fixed by the anchoring contribution of *must* to the present, actual world, the effect of the alternatives to reduce the domain of background uncertainty to gaps in the witnessed set of propositions of the speaker. Against this background, the speaker asserts her choice of alternatives, presumably based on her best indirect evidence and inferential powers. This seems like the right meaning for epistemic *must* in this context. The contextual opening for different ‘Grounds’ for making the choice is what gives epistemic modality its evidential flavour. The fact that the speaker nevertheless has grounds for choosing p over $\neg p$ reflects the fact that they have good evidence, or have inferred based on other information that it is correct. But the existence of the set of uncertainties ensures that this meaning will never be equivalent to a plain assertion.

The modal contribution of *must* in its epistemic use should be essentially the same as the denotation for circumstantial *must* except for the fact that there is no intermediate perspectival situation s' . The CHOICE governs alternatives for the topic of d, by assumption the SPEAKER. Here, f for the depicted event is temporally anchored to the here and the now, and the modal asserts therefore that the actuality of the situation in question is the Speaker’s only choice given their knowledge. But what are the alternatives here? The world and time parameters of f in this case are not ‘up-for-grabs’, since by hypothesis they have been specified by the anchoring contribution of *must*. But this is not surprising, since our intuition here is that epistemic modality is not about circumstances, and by hypothesis it does not traffic in spatiotemporal properties.

Of course, in principle, the speaker has a completely open choice of what to assert about the world. She could choose to say *I am hungry* or *It is snowing like crazy outside*, or she could choose to say *Jane must be in Edinburgh*. The alternatives open to the speaker are in principle endless, and this cannot be what the CHOICE predicate of the modal meaning is operating over. It seems to me that the ‘alternatives’ here must be the simple assertoric options related to the Question Under Discussion (QUD) related to the discourse.

I express the meaning of epistemic *must* more fully then, as in (27).

(27) $[[\text{TP}_{ep-must}]] = \lambda d \exists f \exists e [\text{State}(e) \wedge \wedge f(d)(e) \wedge \perp u \perp(e) \wedge \text{INDEX}(f) \wedge e \text{ is the ONLY assertoric CHOICE for the speaker of } d.]$

Where the LIVE-ALTERNATIVES for the speaker are the different assertions possible given the discourse Question Under Discussion.

We can see that the epistemic modal force of a modal like *must* can indeed vary drastically given the discourse context, even when confined as it is to making an assertion about the current world and time. Consider the following mini-dialogues.

(28) A: Is John in his office?
B: Yes, he must be.

(29) A: Who is in the office now?
B: Mary must be. She always gets there by 8,

(30) A: Where is Mary?
B: She must be in the office.

Thus, when it comes to epistemic modality, the alternatives we are operating over are plausibly the same kinds of Roothian alternatives that are necessary for the construction of focus meanings. I consider this to be positive aspect of the journey to reconfigure modal semantics to be sensitive to syntactic structure in the sense that it employs semantic notions that we already know to be independently necessary in the description of natural language meanings.

If we look at the denotation for the TP headed by *must*, it is just like the present tense except for the addition of some extra content related to the notion of speaker CHOICE among alternatives. One conservative position would be simply to say that this content comes from the lexical presuppositional information contributed by the specific modal and that it does not in fact stem from any higher structural position in the clause. If this were true, then epistemic modals would simply be versions of the modal which lexicalize only T, leaving the Asp feature unassociated (something allowed in principle in the spell-out system I have been assuming). However, the presuppositional content in this case does not go away, but applies instead to the utterance variable as situational perspective. There are some reasons to think however that as a systematic synchronic possibility of this system it might seriously overgenerate. After all, not *all* circumstantial modals have an epistemic counterpart in languages, although it is a common phenomenon. I will assume that the epistemic modal is one that lexicalizes only T (suitably generalized) and that its relation to the lower circumstantial version is fed by grammaticalization. Further discussion of this issue is unfortunately beyond the scope of this monograph.

6.2.1 Aktionsart Sensitivity of Epistemic Modals

However, we do have additional evidence that the temporal specification of *must* is being employed at the TP level and does not require the introduction of an intermediate reference variable. A fact noticed in Ramchand (2014a) is that

under its epistemic reading, *must* is confined to stative prejacent. Consider the data in (31) below. While all kinds of stative prejacent including derived states such as progressive ((31-b)) give perfectly good epistemic readings, a dynamic verb phrase such as ‘write that book’ only gives rise to a deontic interpretation either in the active (31-c) or the passive (31-d).

(31) (a) Jane must be in Edinburgh. *epistemic and deontic*
 (b) Jane must be writing her book. *epistemic and deontic*
 (c) Jane must write that book. *deontic only*
 (d) The book must be written. *deontic only*

This pattern is totally expected under the hypothesis that the indexical temporal specification of *must* applies to the first phase situational description itself under the epistemic reading (without the introduction of an intermediate reference situation) and this has the effect of constraining the situational descriptions to be those that can be true ‘at a moment’. Consider again the Dynamic eventive vs. stative meaning postulate proposed in Taylor (1977), repeated here in (32).

(32) **Events vs. States (inspired by Taylor 1977):**
 (i) If α is a *stative* predicate, then $\alpha(x)$ is true at an interval I just in case $\alpha(x)$ is true at all moments within I;
 (ii) If α is an *eventive* predicate, then $\alpha(x)$ is only true at an interval larger than a moment.

We know that the English present tense is special in requiring a stative complement, and now we know further that it carries over to *must*, but only on the epistemic reading when it lexicalizes T and combines directly with the verbal prejacent.⁶⁴

Now we need to turn to the case of an epistemic modal with anaphoric anchoring specification and see if we predict the right meanings. The epistemic modal *might* has an anaphoric and not indexical specification for worlds and time, and is also different from *must* in asserting simple CHOICE, not exclusive CHOICE.

⁶⁴ Note also that this interpretation of the present tense is consistent with both the perfect and the progressive creating derived states, while the passive does not. This is what we have assumed to be the case in all the denotations given so far. One further construction type that we have not considered in any detail is the use of the habitual present tense on English eventive predicates. I will assume that the habitual also requires the construction of a derived state. I think this is a quite plausible possibility on the surface, but working out the details is once again beyond the scope of my remit here.

(33) Jane might be in Edinburgh

Intuitively, what we want for the meaning of Epistemic Might under this kind of system is what is shown in (34)

(34) **Epistemic Might:**

The proposition expressed is one CHOICE for the speaker at a contextually salient world time pair $\langle w, t \rangle$, given the alternatives open to her, consistent with her knowledge.

This means that the speaker has grounds for thinking that the proposition has a chance of being true, although she does not know it directly. But now the proposition that is being entertained is not actually constrained to be one that holds at the speech time. This predicts in particular that epistemic uncertainty with *might* can be asserted of propositions involving situations at any time at all distinct from the real world. As pointed out in Ramchand (2014a), this is in fact the case, contrastingly sharply with what we found for *must*. So, in the examples shown in (35) we see that epistemic readings are possible for states (35-a) and dynamic events in the future (35-b).

(35) (a) Jane might be in Edinburgh. *epistemic*
 (b) Jane might go to the party. *epistemic*

It is important to emphasize here that the difference between (35-b) and (31) cannot be ascribed to epistemicity per se, but also must be related to the different anchoring properties of the two modals in question. Given this, the denotation for an epistemic modal TP involving *might* should look as in (36). As with epistemic *must*, we assume that the alternatives in question arise from the assertoric alternatives generated by the Question Under Discussion.

(36) $[[\text{TP}_{ep-might}]] = \lambda d \exists f \exists e [\text{State}(e) \wedge \wedge f(d)(e) \wedge \perp u \perp(e) \wedge \text{ANAPH}(f) \wedge 'e \text{ is one assertoric CHOICE for the speaker of } d.]$

Where the **LIVE-ALTERNATIVES** for the speaker are the different assertions possible given the discourse Question Under Discussion.

Thus, the epistemic modals in English come in both anaphoric and indexical flavours, and are characterized by the fact that they express either simple or exhaustive CHOICE with respect to the assertoric alternatives open to the speaker when faced with the Question Under Discussion.

I summarize the epistemic modal meanings and their classification, and their selection with respect to negation in the table below.

| Epistemic Modal | \pm Exclusive Choice | Anchoring |
|------------------------------|------------------------|-----------|
| <i>must</i> ^{Pol} | + | INDEX |
| <i>might</i> ^{Pol} | — | ANAPH |
| <i>should</i> ^{Pol} | + | ANAPH |
| <i>Pol could</i> | — | ANAPH |
| <i>will</i> ^{Pol} | + | INDEX |
| <i>would</i> ^{Pol} | + | ANAPH |
| <i>–Pol can</i> | — | INDEX |

The observant reader will notice that *may* is absent from this list. Permission *may* was listed as *Pol May* in the circumstantial modal list, but there is a question about how to classify the version of *may* found in (37).

(37) (a) If things go according to plan, John may well go that party after all.
 (b) John may not end up passing the exam.

There are two relevant things to note here. One is that *may* is the only modal in the English inventory that is not consistent in its choice of scope with respect to polarity negation. The second is that the use of *may* in the above sentences is very similar to the use of future *will*, which we argued to be a pure circumstantial. In fact, *may* looks like the simple CHOICE version of EXCLUSIVE CHOICE *will* in these contexts.

(38) (a) If things go according to plan, John will maybe go to that party after all.
 (b) John will maybe not end up passing the exam.

The parallelism suggests that if *will* is a prediction circumstantial, then so is *may*. Classifying this *may* as a prediction circumstantial means that there are two circumstantial *mays* in English. We will call permission *may May*₁, and prediction *may May*₂ and assume that we are dealing with two lexical items here. Once there are two lexical items here, *may* is no longer a counterexample to the generalization that modals are specified for selection with respect to negation. The decision to classify *May*₂ as a prediction circumstantial also makes sense of the fact that this reading is not confined to stative prejacent, even though *may* in all its uses has an indexical specification.

This completes the discussion of the general proposal for epistemic meanings. In the following subsections, I look at further extensions and predictions of the model.

6.2.2 Modals Embedding the Perfect

In this section, I look at what happens when a modal like *might* or *must* actually embed the perfect in English. Since this monograph is about auxiliaries and ordering in English, we need to examine each auxiliary not just on its own but in concert with others where such combinations are possible. In fact the interpretation of epistemic modals embedding the perfect produces has been the topic of much recent work in semantics because of the interesting different readings that arise (Condoravdi 2002, Stowell 2004, Demirdache and Uribe-Etxebarria 2008). Consider the sentence in (39), which is claimed to have at least two, and possibly three different readings.

(39) John could have won the race.

- (a) ...let's go and find out. ('Past' Epistemic reading)
- (b) ...but he didn't in the end. (Counterfactual reading)
- (c) ... (still) at that point. (Backshifted or metaphysical reading)

The question is whether the compositional system proposed in this work can generate these readings and their restrictions. In the literature, the PAST EPISTEMIC reading in (39-a) is treated by assuming that the PAST operator contributed by the perfect auxiliary takes scope over the modal. However, I have already committed myself to analysis of the perfect auxiliary in which it introduces an intermediate situational variable related to the lower situation by a relation of inferential prediction. Consider the sentence below with *must* embedding the perfect.

(40) John must have won the race.

The representation built by the perfect auxiliary is a stative situation with inferential force, not a semantic PAST operator. In (40), therefore, by assumption, the meaning we generate is that there is a current situation s' which gives evidence for the event of 'John winning the race'. The situation s' is current because we have assumed that epistemic must contributes a 'present' tense temporal relation to the verbal extended projection. The epistemic modal meaning of *must* now attaches to assert that s' is epistemically forced on the speaker. Thus, it seems that with the current system we already generate the epistemic uncertainty reading for (40), which is good. And we do so without assuming anything special or different about the perfect auxiliary in this construction.

What about the counterfactual, or backshifted ‘metaphysical’ reading of (40), how do we generate that? Well, it is here that the data become interesting because in fact for (40) no such readings exist. This is in contrast to (39) above where those readings naturally emerge. The generalization in fact is that modals with an indexical specification do not give rise to the other two readings at all, and this should give us a clue to the correct analysis of those forms. Below, I show the epistemic indexical *cannot* and *may* embedding the perfect. The reader can check that no counterfactual or backshifted reading for those sentences exist.

(41) (a) John can’t have won the race.
 (b) John may have won the race.

So we see in fact that scope reversal involving a *PAST* operator overgenerates empirically. The counterfactual possibilities are shown by the very modals that exhibit *anaphoric* anchoring behaviour in allowing embedding under a past tense operator.

(42) (a) John could have won the race. *Counterfactual/backshifting possible*
 (b) John might have won the race. *Counterfactual/backshifting possible*
 (c) John should have won the race. *Counterfactual/backshifting possible*

One would hope that the difference between anaphoric and indexical anchoring for the perspective situation would also be able to account for the existence of the counterfactual reading.

We have seen that the epistemic (‘past’) reading is already straightforwardly accounted for under the present system where *PerfP* denotes a derived state which ‘gives evidence for the previous existence of the dynamic eventive situation’. To get the epistemic uncertainty reading for the anaphoric modal anchoring we need only assume that the temporal moment can indeed accidentally overlap with the utterance time, but it is not identified with that single moment. This means that all three modals in (42) can build ‘past’ epistemic readings by asserting that the *s'* inferential state built by the perfect holds at some time overlapping the time of utterance.

However, we need to build the other two readings as well, using the anaphoric flexibility of the above type of modal. In the discussion, I will refer to the reference situation introduced by the perfect as *s'*. This is the situation that gives evidence for the past event of ‘winning’. It is *this* situation that is being judged

as epistemically possible using the epistemic modals *might* and *could*, but crucially under the epistemic reading, it is s' that is anchored using the modal's anchoring information. So, while with epistemic modal *must*, s' had to be current, in the case of the other two modals, the spatiotemporal location s' can be anaphorically resolved.

Thus, *John might have won* says that an s' situation giving evidence for John winning is an assertoric choice for the speaker at some possible world and time as determined by the discourse. Since *might* is not indexical, we can assume that the choice of world is not confined to the actual one, and that any salient enough, and even hypothetical world, time pair can be chosen. If we assume this, then if I we identify the location of s' situation anaphorically to some non-actual $\langle w, * \rangle$ pair, then the situation that gives evidence for 'John win the race' could be located now, but at *some non-actual world*. This will give the straightforward counterfactual reading. If the location of s' is identified as the hypothetical future from a particular point in the past (as in *At that point, he could have still won the race*), then what we get is the metaphysical reading. It is the flexibility of the anchoring properties of the non-indexical modals that allow this kind of contextual sensitivity, and give what looks like a multiplicity of distinct readings. But there seems to be no necessity in this kind of model to describe this in terms of scope reversal. Crucially, none of this possibilities for anchoring the world, time pair of the situation introduced by the perfect is possible for the indexically anchored modals.

Note that we also predict the possibility of deontic modals embedding the perfect, since they inhabit the same zone. But here, the reading we predict is that of the inferential perfect state s' being projected as a possibility or exclusive possibility for the agent. The forward projection of the prejacent event is what we have already built into the semantics of circumstantial modality via the idea of circumstantial **LIVE-ALTERNATIVES**. It is easy to see that these combinations are also licit and get the predicted meaning (although in certain cases some contextualizing is needed, as with all deontic modulations of 'statives').

(43) (a) John must have his homework done by noon.
(b) The applicant may have completed her degree at the time of application.

This ends the discussion of epistemic modality per se. I have proposed that epistemic modals are anchoring elements that unlike the perfect and circumstantial modals, do *not* introduce a reference or perspectival situation, but express the semantics of choice parasitic on the utterance situation itself and the assertoric choices of the speaker. This, intuitively, is the same conclusion

reached by Hacquard in her own seminal work on the differences between circumstantial and epistemic modality. The differences are implementational: the present system involves reification of the utterance situation in shape of the situational variable *d* introduced already at the edge of the first phase. Other major differences revolve around the reconceiving of the general nature of modal meanings as involving CHOICE within LIVE-ALTERNATIVES and tying the nature of those alternatives to the denotation of the prejacent. In this sense, the connection to the compositional system proposed here applies to the analogue of the modal base in addition to the analogue of the perspective situation.

6.3 Evidence for Choice among LIVE-ALTERNATIVES as Basic to Modal Meaning

In this section, I stand back a moment and assess what the shift away from quantificational meanings for modals has done and whether such a drastic move has any further desirable consequences. I am aware that given the pedigree of quantificational analyses of modal meanings, there needs to be quite a number of payoffs if the alternative is to be entertained. The first payoff is that dynamic, circumstantial and epistemic modal meanings under this account can be related straightforwardly to the three domains of the clause that seem to be semantically necessary for independent reasons. The second is that the notion of ‘live alternatives’ makes sense of some traditional longstanding puzzles concerning the weakness of universal modal meanings, and in the interaction of existential modals with overt disjunction. I take these points in turn.

6.3.1 Relation to Dynamic Modal Interpretations

Under a quantificational account of modal meanings, the difference between the lexical modals and the auxiliary modals, reflects a lexical vs. functional distinction in the grammar. The circumstantial and epistemic modals are uniformly argued to be quantificational, while the lexical/dynamic modals seem to have their meanings in the domain of lexical semantics. As we have seen, it is not a given that an individual modal has to have a corresponding lexical alternant, so maybe it is all right if the two types are analysed quite differently. However, it is true that at least in English, a lexical modal meaning does always coexist with a quantificational version. Let us see how we would treat the dynamic modal *can* in English under the present account.

Recall that I have argued in the first half of this chapter that the dynamic modals show all the syntactic hallmarks of low first merge attachment within the lowest event structure domain. This is because they affect the argument roles of the predicate that they combine with and they have the possibility of scoping under a quantificational subject. Since the modal is in the domain

of D_μ where event properties are confined to those that are abstractions over space and time, whatever we say about John's alternatives here must not be dependent on any actual swimming events either before now or in the future. I am forced to conclude that the notion of ability or the notion of disposition is a primitive event property, which can compose with other event properties to create 'the ability to V' and 'the disposition to V' respectively. I suspect that the English habitual is in fact the default specification of the latter meaning.

(44) (a) John can swim. *Dynamic*
 'John possesses the property of having <*Johnswim*> in his abilities to put in train'
 (b) John swims *Habitual*
 'John possesses the property of having <*Johnswim*> in his disposition to put in train.'

It is tempting to think of the causal relationships among subevents as the analogue of the flow of time in the domain of particulars. Basically ability in the force dynamical domain to effect a change is paralleled by circumstantial facilitation in the spatiotemporal domain. The relation to higher modal meanings is that it is these primitive cognitive concepts that are reused in the higher domains via metaphorization to encode more abstract situational versions of these basic meanings. The metaphor involves the relativization of the notion of potentiality and disposition to situational live options, or epistemic options.

Whether or not the dynamic version of the modal *can* in English is the very same lexical item in the grammar of a native speaker or not, the location of the meaning of modality in the domain of 'live options' is a classification that starts to make sense of the reasons such polysemies or historically derived homonimies arise in the first place.

Notice that when derivational morphemes such as *-er* or *able* apply to verbal root symbols, the meanings generated have precisely the kind of pseudointensionality that we have come to expect from the first phase. They are further evidence that the basic meanings of potential and disposition are available at the level of lexical concept formation which is crucially abstracted away from actual real world instantiations.

(45) John is a swimmer/John is a smoker *Disposition*
 This avocado is edible/This movie is unbearable. *Potential*

6.3.2 The Semantic 'weakness' of Universal Modals.

There is an intuition among speakers that making a modal statement is in fact weaker than a simple assertion of fact. However, the formal semantics of

modals like *must* as involving universal quantification over possible worlds makes it seem like a *stronger* reading than the simple assertion of factuality in the actual world. The puzzle has recently come to the fore again through the work of von Fintel and Gillies (2010) who dub the intuition about the ‘weakness’ of *must* The Mantra. Consider the following two utterances.

(46) (a) It is raining outside.
 (b) It must be raining outside.

A speaker in full possession of the facts would be uncooperative in the extreme not to utter (46-a) if it were true. The sentence in (46-b) feels like something of a hedge. However, the classical account in terms of possible worlds involves universal quantification over all worlds in the modal base, including presumably the actual one. Thus it seems as if (46-b) should be saying something that is *even stronger* than (46-a), contrary to our intuitions

One way to fix this is to weaken the modal meaning somehow, for example by saying that the possible worlds in the modal base here contain a hedge to the effect that the real world might *not* be one of the worlds of the modal base consistent with my knowledge. This is the view taken overwhelmingly in the literature starting with Karttunen (1972) who first discussed the problem. See also (kratzer etc.).

Despite the intuition that *must* is somehow ‘weaker’ than plain assertion, von Fintel and Gillies (2010) nevertheless come out in support of the strong view, maintaining that *must* indeed represents universal quantification over a realistic modal base, and that *Must* ϕ really does entail ϕ , and is strong in this sense. The illusion of weakness, they claim arises because of a presupposition that *must* carries concerning the nature of the evidence that allows the speaker to assert that ϕ is true all possible worlds. Essentially, *must* presupposes that the evidence is not direct or perceptual, and it is this evidential flavour that contributes to the pragmatic effect that the speaker is saying something weaker than a plain assertion.

It should be clear that the present proposal for *must* does not have the property of being *stronger* than the simple assertion. In fact, a modalized sentence and a simple assertion are not in any kind of entailment relation in either direction, so no scale of strength can be set up. The proposal for *must* however, does include the idea of LIVE-ALTERNATIVES. This is only well defined in the context of some sort of factual uncertainty, where certain aspects of the situation or proposition are technically undecided. I have claimed that this is always the case when there is an absence of direct witnessing of a situation. From this point of view, the denotation proposed is a direct translation of the intuition that modal statements, even those with the so-called universal modals

must, are always statements made from a backdrop of factual uncertainty. This in fact is built in to their meaning under the current proposal.

**6.3.3 The Puzzle of Deontic Modality and its interaction with Disjunction.
(The Paradox of Free Choice Disjunction (Ross 1941))**

Another semantic puzzle current in the literature on modals is relevant to the current proposal. This is the problem of deontic modality and its interaction with disjunction. Deontic logics are well known for generating paradoxes when compared to the alethic inferential systems they are analogues of (see McNamara 2014 for discussion). This particular problem is well known in the literature and has been taken up recently in Zimmerman (2000) and Aloni (2007), to whom I owe this particular presentation of the problem.

Consider the following sentences of English.

- (47) (a) Vincent is in Paris or in London.
- (b) \Rightarrow Vincent is in Paris or Vincent is in London.
- (48) (a) Vincent may be in Paris or in London.
- (b) \Rightarrow Vincent may be in Paris or Vincent may be in London.
- (c) \Rightarrow Vincent may be in Paris and Vincent may be in London.

The problem is that while (47-a) entails (47-b) in the expected way, and that (48-a) on one reading entails (48-b), it is also true that under a narrow scope reading for the disjunction, (48-a) entails (48-c) !

It seems as if we want the following to be true for permissible (PE):

- (49) $PE(p \vee q) \Rightarrow PE(p) \wedge PE(q)$

But (49) is not a theorem of standard deontic logics, and if we added it as an axiom for PE, then we would end up with an unacceptable system. Because $PE(p) \rightarrow PE(p \vee q)$ is already a theorem in the system, combining that with (49) would have the unwelcome result that if anything at all is permitted, then everything is.

So, on the one hand, we want to preserve a standard logic for disjunction and a standard deontic logic. But we cannot have them both and still generate the entailment in (48-c).

The solution in Zimmerman (2000) is to modify the semantics for disjunction. The solution in Aloni (2007) is to modify the semantics for modals.

Note that the problematic entailment does not arise for universal modals like *must*.

(50) (a) Vincent must be in Paris or in London. $\neg \Rightarrow$
 (b) Vincent must be in Paris and Vincent must be in London.

In the present system, none of the standard paradoxes arise because deontic modality is not characterized in terms of the analogy to alethic modality. Rather, what is proposed here is a notion of ‘choice’ as part of the central *lexical* content of a modal predicate itself. In order to address the problematic entailment, we need only notice that the natural language disjunction *or* is interpreted either as wide scope over the whole proposition to deliver the (b) entailment in (48), or it has low scope under the ‘choice’ predicate. What is then expressed is that ‘being in Paris or London’ is a possible choice for Vincent under the possible circumstantial alternatives. I assume that this means that *or* lies within in the Evt domain on this reading, we need to construct a *single* event property corresponding to ‘being in Paris or London’. All we need to assume here is that a composite event property of this sort has the following characteristic:

(51) **Disjoint Properties and Event Instantiation:**

Any event that instantiates property ‘A or B’ must always have mereological subparts corresponding to an event instantiating A and an event instantiating B.

The reason (51) holds is that because the property is an essential not contingent property, a feature due to our assumptions about the lowest event essences zone, the two disjuncts must be essentially present in all true instantiations of that property. This means that any deontic modal statement that offers the pivot a CHOICE consisting of the situation instantiating the complex property ‘A or B’, will automatically have a situational choice that contains situations instantiating A and B as mereological subparts. This entails that A is a CHOICE for the pivot *and* B is a CHOICE for the pivot.

6.4 Conclusion

This chapter has made a set of specific proposals for how modal and temporal interpretation is integrated into the clause. I have argued that temporal modification is only possible at the level of the second phase where we have the semantic type of properties of spatio-temporal event properties. This is the zone where the spatiotemporal properties of s_0 can be further specified and modified, and where ultimately the anchoring information is expressed. The anchoring information is, specifically, identity or non-identity with some parameter of the utterance situation, d . I argued that present tense in English

expresses indexical anchoring in this sense (identity with the temporal moment of the utterance situation), and past tense expresses non-identity, thus opening up the temporal interval of the described situation to anaphoric resolution. Within modals, we found that they too expressed either indexical or anaphoric anchoring properties as part of their lexical specification. We saw that by and large, the anaphoric modals are the ones in English that still bear moribund ‘past’ tense morphology, probably not accidentally so. This classification is independent of the types of domains that the modals can operate over and independent of whether they are ‘universal’ or “existential” in traditional terms— all modals in English are ‘finite’ and have some anchoring specification for the perspectival situation that they introduce.

With regard to modal meanings per se, we pursued an approach by which the interpretation of the modal could be made to be sensitive to the denotational properties of its complement. I also argued against a quantificational approach to modal meanings, based in part on the unexpected interactions with negation and on the desire to unify the lexical content of modal meaning across zones to as great an extent as possible, and allow the final meaning to be fed directly by the denotation of the prejacent. For this reason, I proposed a view of modality which centres on the notion of *Live-Alternatives*, and CHOICES for an individual within those up-for-grabs alternatives. Under this conception of things, the so-called ‘existential modals’ correspond to a simple assertion that something is ‘a’ choice for the individual, while the ‘universal modals’ correspond the assertion that something is the *only* or exclusive choice for that individual. This general lexical meaning, was argued to be abstract enough to form the basis for particular modal meanings in all three different domains.

Figure 6.4

Relativization of Modal Interpretation

| Zone | Choice Pivot | Source of Uncertainty | First-Merge |
|-----------------------|-------------------|--------------------------------------|-------------|
| Conceptual (Dyn) | Actor | Inherent Causal properties of Actor | Evt |
| Spatiotemporal (Circ) | Situational Topic | Undecidedness of Future Circumstance | Asp |
| Assertoric (Epist) | Speaker | Lack of Complete Knowledge | T |

To summarize my proposal here about the point of attachment of modals, I have assumed that they combine semantically with a projections in the zone which denotes properties of elements of D_μ (the dynamic modals), projections which denote properties of spatiotemporal event properties (circumstantial modals), and projections which denote properties of the utterance (episodic modals). In addition, many modals select directly for PolP. I have as-

sumed that selection for PolP has to be stated on an item by item basis, but that the selection for negation is maintained across polysemous uses even when these straddle zones.

The following phrase structure trees schematically show the possibilities for attachment for epistemic and circumstantial modals respectively. In the diagrams, spans are represented by dotted lines, the @ sign is the Brody-an dia-critic representing the position of linearization for span.

Figure 6.5
Epistemic Modal

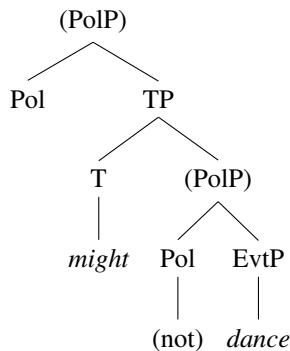
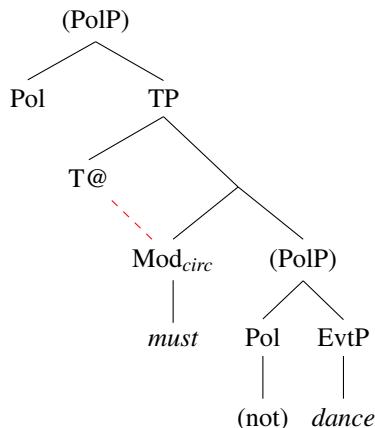


Figure 6.6
Circumstantial Modal



We have now come to the top of the auxiliation ladder and given an explicit treatment of anchoring elements like tense and epistemic modality. The elements that participate in rigid auxiliary ordering in English have now all been

given semantic and syntactic specifications. This has been done in the context of a semantically zoned verbal extended projection, presumed to be universal, and with a view to unifying as far as possible the denotations of the individual pieces. The aim has been to construct a system in which (a) the typological generalizations of meaning compositions across language fall out as a natural consequence of the zoning and (b) the actual ordering of auxiliaries in English does not need to be stipulated by template. In the final chapter I summarize the overall proposal and assess how it meets the goals set out in the beginning of this monograph.

7 Summary and Future Prospects

In this final chapter, I summarize the different proposals made during the course of the book, and then take another look at the sentence in (1),

(1) Vidar might have been being chased.

to see how the system proposed delivers the ordering that we find. Finally, I will discuss extensions and implications of the architecture proposed, and lay out an agenda for further research.

7.1 Architecture and Semantic Zones

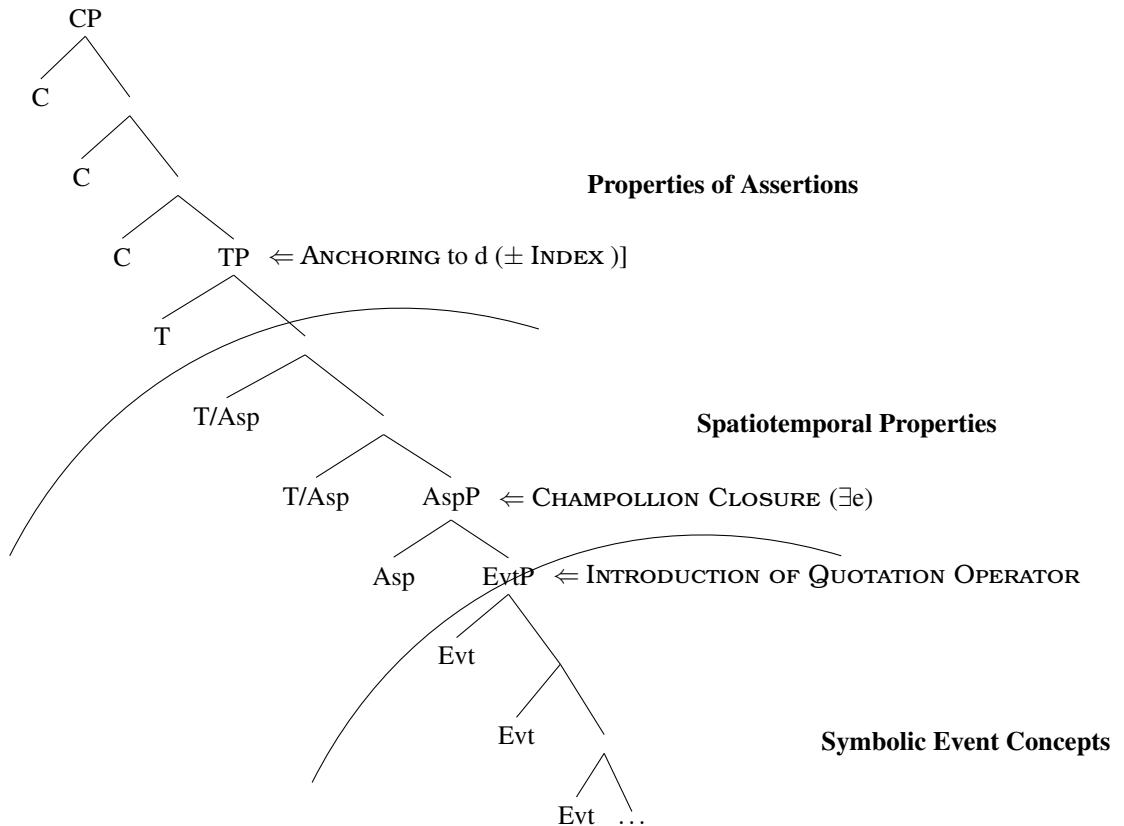
The overall architectural claim of the book has been that the syntax-semantics mapping corresponding to the extended verbal projection delivers denotations in three distinct domains, corresponding to the three distinct hierarchically ordered domains we have robust evidence for in syntax. Syntactic research tends to stipulate the $V < T < C$ hierarchy templatically, but if this work is on the right track, then that hierarchy is mirrored by a hierarchy of semantic sorts. This in turn raises the possibility that the hierarchy of semantic sorts may be derived from something else. My suspicion is that the ‘something else’ is kind of developmental cognitive prioritizing, as well as third factor considerations related to the reusability of symbols in declarative memory. It is not the purpose of this book to argue for a specific explanation in these terms. I have only been at pains to construct a system in which such questions can be asked and tested, where the primes of the semantic ontology are more commensurate with the primes and basic elements being discovered in psycholinguistic and neurolinguistic investigation. If we are to make any progress at all on the question of how much of the linguistic system is determined by constraints imposed by the properties of mind/brain, then semantic ontology needs to be reconfigured from a less symbolic and more algorithmic perspective (to use the terminology of Marr).

The zones I have proposed are the following. First, we have a symbolic, conceptual zone where memorized elements of the system are deployed and composed. The idea here was that these lexical items are stored as $\langle \text{phon}, \text{syn}, \text{sem} \rangle$ triples and that the semantic part of this triple is confined by definition to partial event properties that are independent of temporal and worldly information. The first zone therefore is the zone where elements of D_μ are composed. The second zone, inaugurated by the merge of an explicit quotational operator introducing the event variable for the utterance, d , is the domain of spatiotemporal event properties.⁶⁵ The idea is that spatiotemporal properties are only statable in a context, where there is an anchoring eventuality, or *Origo*, to construct such meanings. In terms of implementation, I close the event variable at the edge of little vP (at the level of what I have called *Evt*), and the second phase becomes essentially properties not of events, but of event *properties*.⁶⁶ Finally, once the outermost f property variable is existentially bound at T , we are left with properties of d , or properties of assertions. This is the final zone of the clause and was necessary to account for epistemic modal interpretations. I also assume that this is the zone of the clause where speaker oriented meanings reside more generally. The zones proposed and their correspondence to the standard syntactic labels for the hierarchy are shown in figure 7.1.

⁶⁵ Recall, this was a move inspired by the semantic work on demonstrations (Davidson 2015 and Henderson 2015) generalized to purely symbolic and non-ideophonic lexical items.

⁶⁶ This move is also made in the work by Lucas Champollion for reasons involving the interaction of the event variable with quantification more generally (Champollion 2015). Regularizing such interactions is part of the payoff here as well, but now it has a more verb phrase internal set of reasons as well—the deployment of the quotational operator and the construction of the spatiotemporal property type.

Figure 7.1
Semantics Zones and Syntactic Domains



The zones proposed have certain clear diagnostic properties. One of the central points made early in this work is that the lowest, symbolic, zone is characterized by meanings that have often been taken to be intensional or modal. My argument has been that all kinds of sub-lexical modality in this sense are really cases of 'pseudo-intensionality' and that they stem from the fact that the symbolic zone manipulates event property abstractions. This zone has its own primitive relations related to identification and subevental relatedness, but does not require the machinery of possible worlds. In the second domain, we find temporal and world parameter manipulations where the full tool box of possible worlds and times is available. In this domain, we also found the possibility of introducing a constrained set of reference situations (specifically in the case of the perfect and of circumstantial modality). I assume that the introduction of a reference situation is constrained by the fact that it needs to bear the kind

of relation to the inner event that will allow specification of *its* spatiotemporal location to give spatiotemporal location to the dependent event. Finally, I have assumed that at the edge of the spatio-temporal zone, we find the locus of anchoring information that explicitly relates the situation to the contextual parameters (I have labelled this T in the phrase structure diagrams, but it may be more suitable in the end to label it INFL following Ritter and Wiltschko 2009). The final zone contains speaker oriented adverbs, metalinguistic negation, and possibly other things like evidentiality markers. The following diagrams summarize, and locate the relevant verbal formatives, as proposed in the previous chapters.

Box 7.1
Symbolic Event Concepts Zone

Verb root, participle in *en/ed*
-ing
Manner and instrumental adverbs
Dummy *be*, Dynamic modality
'constituent' negation.

Box 7.2
Spatiotemporal Properties Zone

Clausal polarity
Circumstantial modality
Perfect *have*
Temporal adverbs

Box 7.3
Properties of Assertions Zone

'High negation'
Epistemic modality
Speaker oriented adverbs.

7.2 Insertion: Lexical vs. Functional Items

In capturing the connection to syntax, I used a view of the lexicon that involved spell-out in the form of spans (Williams 2003, Bye and Svenonius 2012, Adger et al. 2009, Adger 2010) where vocabulary items are specified with a list of category features from the universal spine, and where they spell out spans corresponding to those features. Vocabulary items cannot spell out non-contiguous spans, nor can they spell out structures for which they bear no feature. In constrained circumstances, I do allow the underassociation of category features, but I do not assume that underassociation is freely available. This implementation is designed to transparently enforce the kind of mapping between lexical spellout, morphological structure and syntactic hierarchy which captures the facts traditionally captured by the Mirror Principle (Baker 1985). A transparent system like this is necessary if the zonal ontology is going to have any effects at all on morphological patterns and typological word order patterns. However, this has not been a monograph about theories of lexical insertion, and I leave it open that an equivalently mirror theoretic respecting alternative could also deliver the correct results here. Crucially, in trying to account for order with zones, I have tried to avoid morphology internal devices that might threaten the mapping between morphology and syntax.

In frameworks such as Distributed Morphology (DM), the root is the receptacle for conceptual content but is devoid of syntactic information. The functional structure of the clause is both syntactic *and* the locus of structural semantic information. In my system too, I assume that the nodes of the syntactic tree are associated with formal semantic denotations which represent the structural semantics of the clause. However, these abstract functional meanings need to be supplemented, or clothed with the conceptual content associated with roots, or contentful lexical items.

However, the classical DM model differs in certain ways from the system I have been working with in this book, and that I have advocated more generally in earlier work (Ramchand 2008, Ramchand 2014b). One big difference is that the distinction between conceptual content and structural semantic content is not serial as in the DM architecture, but parallel. The contentful lexical item, moreover, is not devoid of syntactic information, but is quite explicitly a triple containing syntactic, as well as phonological and semantic information. The way this is to be thought of is that the syntactic information provides the link between the LIs conceptual and phonological properties, and governs its

conditions of deployment.⁶⁷ The idea is then that the conceptual content of the LI unifies with the structural semantic content provided by the syntactically labelled heads. I have specified in some detail what kinds of relations and structural semantics I think belong in the first phase, and this is not the place to rehearse them again (they involve causation, and property predication for example). It suffices to say that the structural semantics associated with, say, *init*, or *proc* or *res* are assumed therefore to be *unified* with the conceptual content provided by particular lexical verbs such as *run* vs. *jump*. In other words, *run* provides conceptual content allowing the identification of a process event of ‘running’, and also of the initiation of ‘running’, while *jump* invokes the perceptual and cognitive properties of ‘jumping’ as applied to an initiated process. The process is quite general, and not formally intractable, although it is in tension with our standard means of writing down the denotations of lexical items.

However, there is one respect in which the DM strategy of separating roots from vocabulary items has echoes in the system proposed here. This is the fact that the lowest domain is special. Elements of D_μ are triples, as I have said, and they are ‘deployable’ explicitly by the speaker. It seems like the most sensible way to think of the building of this part of the verbal projection is by the merging of elements of D_μ themselves to build structure. This indeed is what I have assumed. The move to include symbolic elements of the language as part of the domain has the consequence of establishing a real rift between elements of D_μ on the one hand, which are deployable in the first phase, and functional vocabulary items which are *not* members of D_μ and which could be thought of as being late inserted as exponents (contributing phonology and in some cases certain conceptual content in the form of presuppositions). In both cases, we have separation of syntactic contribution and conceptual-phonological matter, but only functional items (i.e. those that are not members of D_μ) could be modelled by late insertion in this system. The D_μ domain is structured, productive and syntactic just as in the constructivist ideal. But it is also similar to the intuition behind roots within DM in that it is encapsulated and early-merged. If one could cite a historical antecedent, the closest in spirit is probably the lexical syntax of Hale and Keyser (2002) in their work through the eighties and nineties.

⁶⁷ In the end, I leave it open that in certain languages and/or for certain items, this syntactic information can be quite underspecified, but it is still a required component of the triple that defines elements of D_μ . The alternative to this kind of syntactic information is the kind of post-syntactic frames of insertion such as proposed in Harley and Noyer 1999, to do the job of subcategorization.

In setting up this system, I have used the tool box that seems to me to be the most transparent for expressing the generalizations that emerge from natural language data, and that is most compatible with my own cherished beliefs. If there were a toolbox that I could have just taken off the shelf to implement the agenda of this book, I would have done so. This was part of the point. However, the architecture I have used has certain properties that are surely contingent and could have been done another way, but the difficulty of capturing the intuition behind event kinds/concepts/essences in a semi-formalized system that makes it possible to examine and extract predictions from has been the most difficult constraint to satisfy. I think it is interesting that this latter challenge has lead me to a system that enforces a stronger distinction between lexical and functional items than I had hitherto assumed, more in line with the position always held by DM. However, because of being conscious of the aim of providing denotations for LIs as polysemous items of representational integrity, I have been forced into a more piece-based, Lego-style approach⁶⁸ to the construction of linguistic representation. This in turn makes it easier to see what predictions this kind of compositional system makes for the storing and processing of LIs in practice in the production and comprehension of natural language propositions.

That said, I summarize next the syntactic and semantic denotations of the core players in the analysis of the English auxiliary system as it has emerged from this study.

7.3 Summary of the Pieces

The *-ing* LI is a member of D_μ and has the following denotation. Its syntax is simply specified as $\langle \text{Evt} \rangle$.

$$(2) \quad \llcorner u_V \text{-ing} \llcorner = \lambda x \lambda e [\text{State}(e) \wedge \text{ID-State}(e, \llcorner u_V \llcorner) \wedge \text{HOLDER}(e) = x]$$

Box 7.4
Identifying State (ID-STATE): Definition

For all event descriptions P , an *Identifying-State* for P , is a stative eventuality that manifests sufficient cognitive/perceptual identifiers of the event property P .

⁶⁸ Thanks to Sandra Ronai (pc) for a suggestion of the metaphor.

The participle in EN/ED is the spellout form corresponding to systematic underattachment of the features $\langle \text{Asp}, \text{Evt}, \text{init}, \dots \rangle$ of the corresponding past tense verb. The underattachment is rooted at the bottom of the span.,

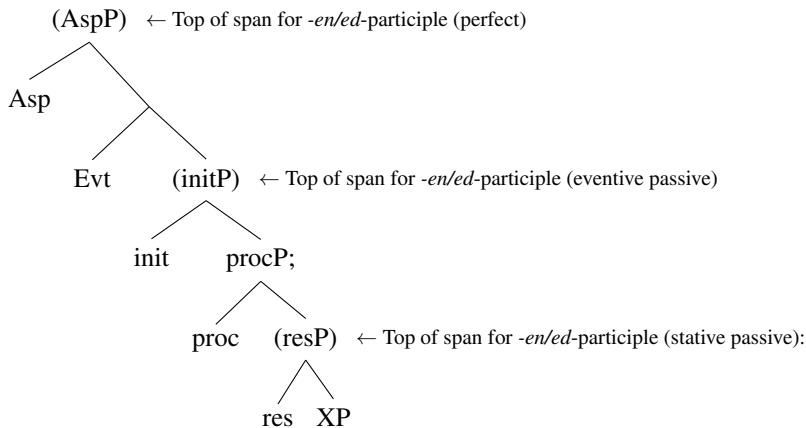


Figure 7.2

Scope of Spell-Out for the -EN/ED-Participle

The following summarizes the blocking facts.

(3) • Attach *-ing* to any complete event structure and fill in with dummy verb *be*:
blocked by stative verbs.

• Spell out resP as *-en/ed* participle and spell out Evt with dummy verb *be*
blocked by adjectives in the case of deadjectival verbs.

• Spell out procP as *-en/ed* participle and spell out Evt with dummy verb *be*:
blocked by unaccusative verbs.

The Blocking Principle for Auxiliation is given in (4)

(4) **Blocking of Auxiliation:**

In cases where a single verbal lexical item generates the same Event description as an Auxiliary structure, expression by means of an auxiliary is blocked.

Have has the syntactic specification <Asp> and has the semantic denotation in (5). It introduces a reference state from which the most embedded event is inferable. The definition of EvidState follows in (6).

$$(5) \quad [[\text{have}]] = \lambda Q \lambda x \lambda f' \exists s' \lambda d \exists f [Q(f) \wedge \wedge \text{State}(s') \wedge \text{HOLDER}(s') = x \wedge \\ 'f\text{-Is-Inferable-from-}s'' \wedge f'(d)(s')]$$

(6) **Evidential State (EVID-STATE)** (Definition)

EVID-STATE FOR $s_0 =_{\text{def}} s'$ iff s' is a stative situation (i.e. which can have a moment as its temporal parameter) which is a salient situation that provides *criterial evidence for the existence of s_0* . The existence of s' always entails the existence of s_0 .

Circumstantial modal meanings are also merged in the second phase, and modify spatiotemporal properties of the embedded event by introducing a perspectival intermediate situation with respect to which the embedded event is said to be a CHOICE (either simple or exclusive) among LIVE-ALTERNATIVES. In the spatiotemporal domain, live alternatives are all fs that locate the embedded event at world time pairs that are not part of a realistic modal base.

$$(7) \quad [[\text{Mod}_{\text{circ}}]] = \lambda Q \lambda x \lambda f' \exists s' \lambda d \exists f [Q(f)(d) \wedge \wedge \text{State}(s') \wedge \text{HOLDER}(s') \\ = x \wedge f = \lambda s \lambda d [s \text{ is located at a world-time pair that is a CHOICE for} \\ \text{the perspectival topic in } s'] \wedge f'(s')(d)]$$

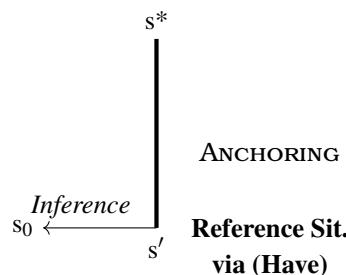
$$(8) \quad [[\text{Mod}_{\text{circ-must}}]] = \lambda Q \lambda x \lambda f' \exists s' \lambda d \exists f [Q(f)(d) \wedge \wedge \text{State}(s') \wedge \text{HOLDER}(s') \\ = x \wedge f = \lambda s \lambda d [s \text{ is located at a world-time pair that is the ONLY CHOICE} \\ \text{for the perspectival topic in } s'] \wedge f'(s')(d)]$$

Box 7.5
Abstract Schema for Modal Denotations

A modal meaning involves the assertion of a CHOICE within a set of ‘live alternatives’ for a Topic individual x in a perspectival situation s' . These alternatives are directly constructed from the constituent that the modal attaches to.

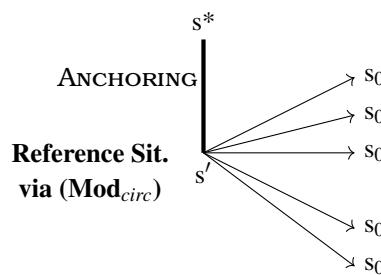
Both the perfect and circumstantial modality involve modification of properties of spatiotemporal properties of events. They have an extremely similar semantic contribution: they introduce a perspectival situation s' that provides an intermediate reference situation between the depicted event and d . This is shown schematically in the following two figures.

Figure 7.3
Schema for The Perfect



In figure 7.4, we see the corresponding picture for the circumstantial modal. The modal auxiliary also introduces an intermediate, or reference situation (or perspectival situation) which is the outer situational variable eventually subject to anchoring. But now the relationship between s' and s_0 is projective instead of inferential.

Figure 7.4
Schema for Circumstantial Modality



Live Alternatives



Comparing the two figures, we can see that it expresses a metaphysical modal base schema (as described in Werner (2006) and Condoravdi (2002)) in which

worlds up to a given time are strictly identical (the actual world) and only diverge in the future of that given time. Within the metaphysical scheme, the perfect asserts the base situation to be in the determined past of the reference situation; the circumstantial modal asserts the base situation to be in the projective nondetermined future of the reference situation.

At the T position of the clausal spine, anchoring to the context takes place. I have argued that there are two basic forms of anchoring. The representations of present and past in English are given in (9) and (10) respectively.

Box 7.6
Types of Anchoring

\forall eventualities e and speech events d, such that $f(d)(e)$, INDEX(f) is true iff the temporal parameter of e is strictly identified with the temporal parameter of d, the speech time.
 \forall eventualities e and speech events d, such that $f(d)(e)$, ANAPH(f) is true iff the temporal parameter of e is resolved *anaphorically*, either by binding from something in the linguistic context, or to some purely discourse contextual topic time or world.

$$(9) \quad [[\text{TP}_{\text{pres}}]] = \lambda d \exists f \exists e [\text{Utterance}(d) \wedge f(e)(d) \wedge \text{INDEX}(f) \wedge \llcorner u \lrcorner(e) \wedge \text{Source}(d) = \text{'Speaker'} \wedge \text{Goal}(d) = \text{'Hearer'}]$$

$$(10) \quad [[\text{TP}_{\text{past}}]] = \lambda d \exists f \exists e [\text{Utterance}(d) \wedge f(e)(d) \wedge \text{ANAPH}(f) \wedge \llcorner u \lrcorner(e) \wedge \text{Source}(d) = \text{'Speaker'} \wedge \text{Goal}(d) = \text{'Hearer'}]$$

All modals in English carry anchoring information. In (11), I give the TP corresponding to circumstantial modal *must* which is anchored indexically, combined with its circumstantial modal meaning.

$$(11) \quad [[\text{TP}_{\text{must}}]] = \lambda d \exists f' \exists s' \exists f \exists s_0 [\text{Utterance}(d) \wedge \text{State}(s') \wedge f(d)(s_0) \wedge \llcorner u \lrcorner(s_0) \wedge \text{INDEX}(f') \wedge f = \lambda s \lambda d [s \text{ is located at a world-time pair that is the ONLY CHOICE for the perspectival topic in } s'] \wedge f'(d)(s')]$$

Finally, epistemic modality occurs when the modal spells out only at T, and applies the CHOICE semantics to d as its perspectival situation. Here the live alternatives come from the speaker's different assertoric choices given their knowledge and the question under discussion.

$$(12) \quad [[\text{TP}_{\text{ep-must}}]] = \lambda d \exists f \exists e [\text{State}(e) \wedge f(d)(e) \wedge \llcorner u \lrcorner(e) \wedge \text{INDEX}(f) \wedge e \text{ is the ONLY assertoric CHOICE for the speaker of } d.]$$

Where the LIVE-ALTERNATIVES for the speaker are the different assertions possible given the discourse Question Under Discussion.

(13) $[[\text{TP}_{ep-might}]] = \lambda d \exists f \exists e [\text{State}(e) \wedge \wedge f(d)(e) \wedge \sqsubset u \sqsubset(e) \wedge \text{ANAPH}(f)$
 $\wedge e \text{ is the ONLY assertoric CHOICE for the speaker of } d.]$
 Where the **LIVE-ALTERNATIVES** for the speaker are the different assertions possible given the discourse Question Under Discussion.

7.4 Auxiliary Ordering Revisited

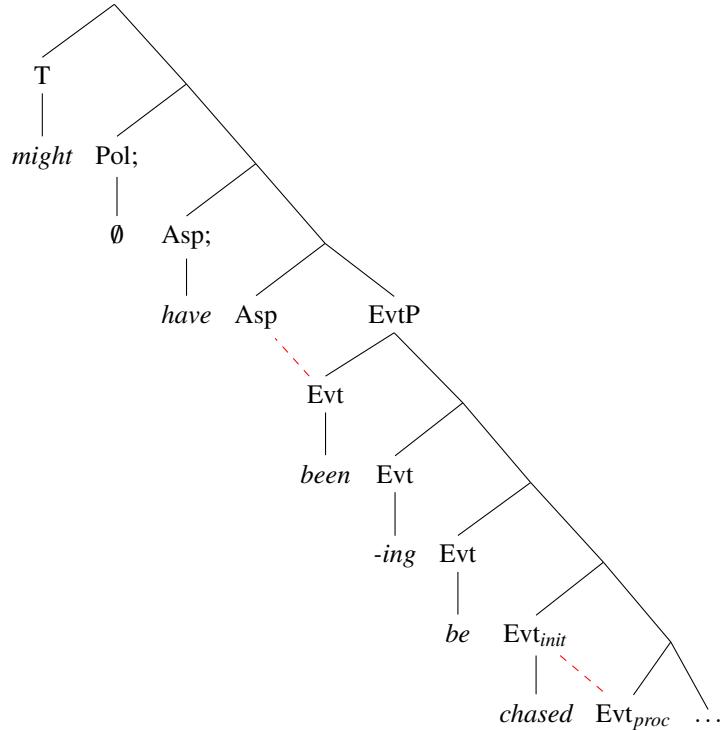
From a basic empirical perspective, the simple fact of auxiliary ordering in English which has been our test case all along now looks like it can be made to follow from zonal properties and ontological classification of denotations rather than by detailed selectional mechanisms for particular morphological endings. The goal has been unified syntactic denotations for all the formatives employed in the auxiliary system, and this, largely, has been achieved. It is time to put all the pieces together, and assess the costs and stipulations that this particular model has required.

As a way of focusing the discussion, let us look at the sentence repeated in (14), with its full complement of auxiliaries in the English system.

(14) Vidar might have been being chased.

The full tree with its spell out by vocabulary items is shown in Figure (14) below.

Figure 7.5
Maximal Auxiliation in English



Chase is a verb with the syntactic specification $\langle \text{Asp}^*, \text{Evt}, \text{init}, \text{proc} \rangle$, and its *en/ed*-participle has the specification $\langle \text{Asp}, \text{Evt}, \text{init}, \text{proc} \rangle$. The participle inserts to spell out *proc* and *init*, and *be* inserts in *Evt* followed by *ing* and then another instance of *be* in the higher *Evt*. Could the ordering have been otherwise? Let us consider the different possibilities.

The progressive could not occur before the passive because the participle involves the spell out of a contiguous span based on the root's specification. Merging *ing* would close off the domain spellable out by the participle. The *ing* could not attach to the participle itself since the participle cannot accept suffixation, which is a requirement of *-ing*. The progressive *-ing* can only attach to the bare root which also spells out the full span (minus Asp^*). Nor would it help to spell out the dummy verb *be* as the participial form. We *can* do it, but it will not make any difference since *be* does not have any other category features other than *Evt* (and Asp^* and *T*) anyway so spelling out as *been* will only affect its ability to enter the situational zone, another *be* would have to be inserted and tensed in any case and the redundant *been* would be blocked. In any

case, it would not give any kind of passivization effect. In fact, as we have seen, we can get *en/ed*-participle above progressive, but it goes on to grow into the *perfect*, since spelling out the participle at this height will inevitably mean that no argument has been removed. Similarly, perfect cannot occur underneath the progressive or passive or dynamic modality since (i) perfect *have* merges in the higher situational domain (by stipulation) and (ii) the effect of non-removal of an argument comes from the very height of the participial span that feeds the perfect and by definition it could not be constructed ‘lower’.

Could dynamic modality occur before progressive? No, this is ruled out by the very same semantic economy rule that prohibits the progressivization of stative verbs. In fact, we already have a ready-made English specific factor which will rule out any kind of modal embedding, which is the fact that in English, modals only have one lexical entry, the one that contains both Asp and T features. There simply are no participial or bare forms for the English modals in the standard dialect.

Turning to the situational zone, could the perfect occur above circumstantial modality instead of below it? I assume that this is possible in principle, but that the language specific fact about English modals as stipulated above prevents it. In languages where modals have non-finite forms, the circumstantial modal can remain in Asp and in principle an epistemic modal could merge to spell out T (as in Swedish or Norwegian). However, in English, the epistemic and circumstantial modals simply cannot cooccur.

7.5 Open Questions and Further Research

7.5.1 The Nominal Domain

The observant reader will have noticed that I have said absolutely nothing so far about verbs. There is a lot of work on the cartography of the nominal extended projection that shows that here too, there is an intriguing typologically robust order to the construction of DPs from base lexical items (Cinque 2004, Zamparelli 2000, Borer 2005, Dékaný 2012, Pfaff 2015a). My assumption here is that nominal projections too are partitioned into a symbolic D_μ domain and a higher domain of instantiation, where I assume both reference, and case reside.

Apart from investigating a similar kind of hypothesis in the nominal domain however, further research must also be specific about how nominal arguments are merged with/integrated with the verbal functional sequence in the course of building up a proposition. In order to gain the advantages of Champollion closure at EvtP for the interaction with quantification more generally, we need to assume crucially that quantified nominal projections are not merged in com-

plete form within the EvtP. I believe that in order to understand the relationship between the two extended projections, we need to adopt a view of phrase structure that involves merging of minimal nominal structures in the lowest minimal part of the verbal extended projection together with higher copies that contain more and more functional information. Thus, I would argue, the part of the argument that is merged in the first phase is actually not a fully fledged referential projection or phase, but the lower portion of the nominal argument, that contributes its conceptual semantics to the build up of the D_μ domain before the deployment operator is merged. Only later on are these arguments given referential status, existentially bound, or quantified. There is some evidence that the direct complements to V should only be N, and not DP (Williams 2003, Sportiche, Svenonius 2004). Taking this result seriously and integrating it into a coherent theory of merge and linearization is entirely possible and is likely to be warranted on independent grounds anyway. One way of implementing the idea would involve the multidominance of the banyan trees proposed in Svenonius (2004). A detailed exposition of the interleaving of nominal and verbal functional sequences to build the proposition is however beyond the scope of this monograph, and will have to be left for future research.

7.5.2 On the Universal vs. the Language Specific

The zones should be universal, since they are designed to account for robust crosslinguistic generalizations in the first place. However, there are a number of features of the auxiliary ordering analysis in English that are clearly language specific.

The presuppositional/conceptual semantics of the language particular lexical items that spell out those zones is of course a matter that is up to each language. In particular, the specific spans spelled out by individual lexical items are specific to the language and the particular vocabulary item. Idiosyncratic facts about English, for example, include the fact that all modals have a T feature and do not have corresponding uninflected entries. The English present tense also turns out to be special and I have speculated that this is one reason why it makes such liberal use of auxiliation in building derived states for anchoring.

In principle from this system, we expect ordering constraints and patterns because of the universal hierarchical ordering of the three semantic zones of the clause. Within this broad expectation, individual languages might impose even stricter and more fine grained orderings because of language specific items. Thus, the fact that there are only three zones in this system does not derive all of the full richness of orderings reported in the cartographic literature. However, language specific selectional facts may sit on top of the more minimalistic

universal spine in this sense (See also Ramchand and Svenonius 2014 and Wiltschko (2014) for discussion of the same general idea).

7.5.3 The Future and the Search for Explanations

The new ontology proposed here offers hope of a more systematic connection with T psycholinguistics and neurolinguistics. Specifically, they are formulated in such as way as to make possible predictions about what we might find in those subfields. As linguists, we are also interested in the idea of how the linguistic system is constrained by the more general properties of mind brain. Language in turn is interesting to those other domains of inquiry because it offers insight into one of the most complicated things that human minds do, namely the processing of hierarchical symbolic structures.

I have hoped in this short monograph to give a kind of proof of concept that a system created with such a radically different ontology can be made to work and meet basic descriptive desiderata. Many details and consequences have remained unexplored. Further investigation will have to be left for future research.

Bibliography

Abney, Steven. 1987. The English noun phrase in its sentential aspect. PhD diss, MIT, Cambridge, Ma..

Abusch, Dorit. 1985. On verbs and time. PhD diss, University of Massachusetts at Amherst.

Adelman, James, Gordon Brown, and José Quesada. 2006. Contextual diversity, not word frequency, determines word naming and lexical decision times. *Psychological Science* 17 (9): 814–823.

Adger, David. 2010. Variability and grammatical architecture. LingBuzz: <http://ling.auf.net/lingBuzz/001176>.

Adger, David, Daniel Harbour, and Laurel Watkins. 2009. *Mirrors and microparametric phrase structure beyond free word order*. Cambridge University Press.

Aelbrecht, Lobke, and Will Harwood. 2012. To be or not to be elided: VP ellipsis revisited. Ms. University of Ghent.

Åfarli, Tor A. 1989. Passive in Norwegian and in English. *Linguistic Inquiry* 20 (1): 101–108.

Åfarli, Tor A. 1992. *The syntax of norwegian passive constructions*. Amsterdam: John Benjamins.

Alexiadou, Artemis, and Elena Anagnostopoulou. 2008. Structuring participles. In *Proceedings of WCCFL*, eds. Charles B. Chang and Hannah J. Haynie, Vol. 26, 33–41. Somerville, MA: Cascadilla Press.

Alexiadou, Artemis, Monika Rathert, and Armin von Stechow. 2003. *Perfect explorations*. Berlin: Mouton de Gruyter.

Aloni, Maria. 2007. Free choice, modals, and imperatives. *Natural Language Semantics* 15: 65–94.

Altshuler, Daniel. 2016. *Events, states and times*. De Gruyter.

Anagnostopoulou, Elena. 2003. Participles and voice. In *Perfect explorations*, eds. Monika Rathert Artemis Alexiadou and Armin von Stechow, 1–36. Mouton de Gruyter.

Austin, J. L. 1950. *Truth: Philosophical papers*. Oxford: Oxford University Press.

Baker, Mark. 1985. The Mirror Principle and morphosyntactic explanation. *Linguistic Inquiry* 16 (3): 373–415.

Baker, Mark C. 2003. *Lexical categories*. Cambridge studies in linguistics. Cambridge: Cambridge University Press.

Baker, Mark C., Kyle Johnson, and Ian Roberts. 1989. Passive arguments raised. *Linguistic Inquiry* 20 (2): 219–251.

Baltin, Mark. 2006. The non-unity of VP-preposing. *Language* 82 (4): 734–766.

Baltin, Mark R. 1989. Heads and projections. In *Alternative conceptions of phrase structure*, eds. Mark R. Baltin and Anthony S. Kroch, 1–16. Chicago: University of Chicago Press.

Barwise, Jon, and John Perry. 1983. *Situations and attitudes*. Cambridge, Ma.: MIT press.

Beaver, David, and Cleo Condoravdi. 2007. On the logic of verbal modification. In *Proceedings of the 16th Amsterdam colloquium*, eds. Paul Dekker Maria Aloni and Floris Roelofsen. Amsterdam, Netherlands: University of Amsterdam.

Bhatt, Rajesh. 2008. Transitivity alternations and verbalization. ms., University of Massachusetts, Amherst.

Bjorkmann, Bronwyn. 2011. Be-ing default. the morphosyntax of auxiliaries. PhD diss, Massachusetts Institute of Technology, Cambridge, MA.

Bobaljik, Jonathan David, and Susi Wurmbrand. 1999. Modals, Raising and A-Reconstruction. ms, University of Connecticut.

Bonomi, Andrea. 1997. The progressive and the structure of events. *Journal of Semantics* 14 (2): 173–205.

Borer, Hagit. 2005. *Structuring sense: An exo-skeletal trilogy*. New York: Oxford University Press.

Borer, Hagit. 2013. *Taking form: Structuring sense volume iii*. Oxford: Oxford University Press.

Bošković, Željko. 2013. Now I'm a Phase, now I'm not a phase. ms., University of Connecticut.

Brennan, Virginia. 1993. Root and epistemic modal auxiliary verbs. PhD diss, University of Massachusetts at Amherst.

Bresnan, Joan. 1982. The passive in lexical theory. In *The mental representation of grammatical relations*, ed. Joan Bresnan, 3–86. Cambridge, Ma.: MIT Press.

Brody, Michael. 2000. Mirror theory: Syntactic representation in perfect syntax. *Linguistic Inquiry* 31 (1): 29–56.

Brown, Roger. 1973. *A first language: The early stages*. London: George Allen & Unwin.

Bruening, Benjamin. 2014. Word formation is syntactic: Adjectival passives in english. *Natural Language & Linguistic Theory* 32: 363–422.

Bye, Patrik, and Peter Svenonius. 2012. Non-concatenative morphology as an epiphenomenon. In *The morphology and phonology of exponence: The state of the art*, ed. Jochen Trommer, 427–495. Oxford: Oxford University Press. ms., University of Tromsø.

Caha, Pavel. 2009. The nanosyntax of case. PhD diss, University of Tromsø.

Carrier, Jill, and Janet H. Randall. 1992. The argument structure and syntactic structure of resultatives. *Linguistic Inquiry* 23 (2): 173–234.

Champollion, Lucas. 2015. The interaction of compositional semantics and event semantics. *Linguistics and Philosophy* 38: 31–66.

Chomsky, Noam. 1957. *Syntactic structures*. 's-Gravenhage: Mouton.

Chomsky, Noam. 1970. Remarks on nominalization. In *Readings in English transformational grammar*, eds. Roderick A. Jacobs and Peter S. Rosenbaum, 184–221. Washington, D.C.: Georgetown University Press.

Chomsky, Noam. 1981. *Lectures on government and binding. Studies in generative grammar* 9. Dordrecht: Foris.

Chomsky, Noam. 1995. Language and nature. *Mind* 104 (413): 1–61.

Cinque, Guglielmo. 1999. *Adverbs and functional heads: A cross-linguistic perspective*. New York: Oxford University Press.

Cinque, Guglielmo. 2004. Deriving Greenberg's Universal 20 and its exceptions. ms. University of Venice.

Cipria, Alicia, and Craige Roberts. 2000. Spanish imperfecto and pretérito: Truth conditions and aktionsart effects in a situation semantics. *Natural Language Semantics* 8 (4): 297–347.

Condoravdi, Cleo. 2002. Temporal interpretation of modals. modals for the present and for the past. In *The construction of meaning*, eds. Stefan Kaufmann David Beaver Luis Casillas and Billy Clark, 59–87. Stanford, CA: CSLI Publications.

Copley, Bridget. 2002. The semantics of the future. PhD diss, Massachusetts Institute of Technology.

Copley, Bridget, and Heidi Harley. 2015. A force-theoretic framework for event structure. *Linguistics and Philosophy* 32 (2): 103–158.

Cormack, Annabel, and Neil Smith. 2002. Modals and negation in english. In *Modality and its interaction with the verbal system*, ed. Sjef Barbiers, 143–173. John Benjamins.

Davidson, Donald. 1967. The logical form of action sentences. In *The logic of decision and action*, ed. Nicholas Rescher. Pittsburgh, Pa: University of Pittsburgh Press.

Davidson, Kathryn. 2015. Quotation, demonstration and iconicity. *Linguistics and Philosophy* 38: 477–520.

Dékaný, Eva. 2012. A profile of the hungarian dp. PhD diss, University of Tromsø.

Demirdache, Hamida, and Myriam Uribe-Etxebarria. 2000. The primitives of temporal relations. In *Step by step. essays on minimalist syntax in honour of Howard Lasnik*, eds. Roger Martin, David Michaels, and Juan Uriagereka, 157–186. Cambridge, Ma.: MIT Press.

Demirdache, Hamida, and Myriam Uribe-Etxebarria. 2008. Scope and anaphora with time arguments: The case of 'perfect modals'. *Lingua* 118 (11): 1790–1815.

Dowty, David R. 1979. *Word meaning and Montague grammar: The semantics of verbs and times in generative semantics and in Montague's PTQ*. Dordrecht: Reidel.

Dryer, Matthew S., and Martin Haspelmath. The world atlas of language structures online.

Eckhardt, Regine. 2012. *Hereby explained: an event-based account of performative utterances*. *Linguistics and Philosophy* 35: 21–55.

Embick, David. 2004. Structure of the resultative participles in english. *Linguistic Inquiry* 35 (3): 355–592.

Embick, David, and Alec Marantz. 2008. Architecture and blocking. *Linguistic Inquiry* 39 (1): 1–53.

Embick, David, and Ralf Noyer. 2001. Movement operations after syntax. *Linguistic Inquiry* 32 (4): 555–595.

Ernst, Thomas. 2002. *The syntax of adjuncts*. Cambridge: Cambridge University Press.

Fine, Kit. 2000. Semantics for the logic of essence. *Journal of Philosophical Logic* 29: 543–584.

Fine, Kit. 2005. Necessity and non-existence. In *Modality and tense: Philosophical papers*, 321–353. Oxford University Press. Chap. 9.

Fine, Kit. 2014. Truthmaker semantics for intuitionist logic. *Journal of Philosophical Logic* 43 (2): 549–577.

Gehrke, Berit. 2015. Adjectival participles, event kind modification: the case of frequency adjectives. *Natural Language and Linguistic Theory* 33: 897–938.

Gehrke, Berit, and Louise McNally. 2015. Distributional modification: the case of frequency adjectives. *Language* 91(4): 837–868.

Giorgi, Alessandra, and Fabio Pianesi. 1997. *Tense and aspect: From semantics to morphosyntax*. New York: Oxford University Press.

Grimm, Scott, and Louise McNally. 2015. The *ing* dynasty: Rebuilding the semantics of nominalizations. In *Proceeding of SALT XXV*, eds. Mary Moroney Sarah D’Antonio and Carol Rose Little, 82–102.

Grimshaw, Jane. 1979. Complement selection and the lexicon. *Linguistic Inquiry*.

Grimshaw, Jane. 1990. *Argument structure. Linguistic inquiry monographs; 18*. Cambridge: MIT Press.

Groenendijk, J., and M. Stokhof. 1975. Modality and conversational information. *Theoretical Linguistics* 2: 61–112.

Hacquard, Valentine. 2006. Aspects of modality. PhD diss, Massachusetts Institute of Technology, Cambridge, MA.

Hale, Ken, and Samuel Jay Keyser. 2002. *Prolegomenon to a theory of argument structure. Linguistic inquiry monograph*. Cambridge, Ma.: MIT Press.

Hale, Kenneth, and Samuel Jay Keyser. 1993a. On argument structure and the lexical expression of syntactic relations. In *The view from building 20: Essays in linguistics in honor of Sylvain Bromberger*, eds. Kenneth Hale and Samuel Jay Keyser. *Current studies in linguistics*, 53–109. Cambridge, Ma.: MIT Press.

Hale, Kenneth, and Samuel Jay Keyser, eds. 1993b. *The view from building 20: Essays in linguistics in honor of Sylvain Bromberger. Current studies in linguistics*. Cambridge, MA: MIT Press.

Halle, Morris, and Alec Marantz. 1993. Distributed morphology and the pieces of inflection. In *The view from building 20: Essays in linguistics in honor of Sylvain Bromberger*, eds. Kenneth Hale and Samuel Jay Keyse, 111–176. Cambridge, Ma.: MIT Press.

Hallman, Peter. 2009a. Instants and intervals in the event/state distinction. ms., UCLA.

Hallman, Peter. 2009b. Proportions in time: Interactions of quantification and aspect. *Natural Language Semantics* 17: 29–61.

Hankamer, Jorge, and Ivan Sag. 1976. Deep and surface anaphora. *Linguistic Inquiry* 7 (3): 391–428.

Harley, Heidi. 2013. External arguments and the mirror principle. on the distinctness of voice and v. *Lingua* 125: 34–57.

Harley, Heidi, and Rolf Noyer. 1999. State of the article: Distributed morphology. *GLOT International* 4.4: 3–9.

Harwood, Will. 2011. There are several positions available: English intermediate subject positions. In *Proceedings of ConSOLE XIX*, 1–25.

Harwood, Will. 2014. Being progressive is just a phase: celebrating the uniqueness of progressive aspect under a phase-based analysis. *Natural Language and Linguistic Theory*.

Harwood, William. 2013. Being progressive is just a phase: Dividing the functional hierarchy. PhD diss, University of Ghent, Ghent, Belgium.

Heim, Irene, and Angelika Kratzer. 1998. *Semantics in generative grammar*. Oxford: Blackwell.

Henderson, Robert. 2015. Pluractional demonstrations. paper presented at SALT 26.

Higginbotham, James T. 2007. Remarks on compositionality. In *Oxford handbook of linguistic interfaces*, eds. Gillian Ramchand and Charles Reiss, 425–444. Oxford University Press.

Homer, Vincent. 2012. Neg raising and positive polarity: The view from modals. *Semantics and Pragmatics*.

Huddleston, R. 1974. Further remarks on the analysis of auxiliaries as main verbs. *Foundations of Language* 11: 215–229.

Iatridou, Sabine. 1990. About agr(p). *Linguistic Inquiry* 21 (4): 551–577.

Iatridou, Sabine. 2000. The grammatical ingredients of counterfactuality. *Linguistic Inquiry* 31 (2): 231–2.

Iatridou, Sabine, and Hedde Zeilstra. 2013. Negation, polarity and deontic modals. *Linguistic Inquiry*.

Inoue, Kyoko. 1979. An analysis of the English present perfect. *Linguistics* 17: 561–589.

Isard, Steven. 1974. What would you have done if... *Theoretical Linguistics* 1: 233–255.

Israel, Michael. 1996. Polarity sensitivity as lexical semantics. *Linguistics and Philosophy* 19: 619–666.

Jackendoff, Ray. 1972. *Semantic interpretation in generative grammar. Current studies in linguistics*. Cambridge, Ma.: MIT Press.

Julien, Marit. 2000. Syntactic heads and word formation: A study of verbal inflection. PhD diss, University of Tromsø.

Julien, Marit. 2002. *Syntactic heads and word formation. Oxford studies in comparative syntax*. New York: Oxford University Press.

Kamp, Hans, and Uwe Reyle. 1993. *From discourse to logic*. Dordrecht: D. Reidel.

Kamp, Hans, Uwe Reyle, and Antje Rossdeutscher. 2016. Perfects as feature shifting operators. to appear, Brill Publishing.

Karttunen, Lauri. 1972. Possible and must. In *Syntax and semantics*, ed. J. Kimball, Vol. 1, 1–20. Academic Press.

Kayne, Richard S. 1994. *The antisymmetry of syntax*. Cambridge, MA.: MIT Press.

Klein, Wolfgang. 1992. The present perfect puzzle. *Language*.

Klein, Wolfgang. 1994. *Time in language*. London and New York: Routledge.

Klinedinst, Nathan. 2012. Intensionality and the progressive, Technical report, UCL, London.

Kratzter, Angelika. 1977. What “must” and “can” must and can mean. *Linguistics and Philosophy* 1 (1): 337–355.

Kratzter, Angelika. 1981. The notional category of modality. In *Words, worlds and contexts*, eds. Hans-Jürgen Eikmeyer and Hans Reiser, 38–74. Berlin/New York: de Gruyter.

Kratzter, Angelika. 1989. An investigation of the lumps of thought. *Linguistics and Philosophy* 12: 607–653.

Kratzter, Angelika. 1991. Conditionals. In *Semantics: An international handbook of contemporary research*, eds. Arnim von Stechow and Dieter Wunderlich. De Gruyter.

Kratzter, Angelika. 1996. Severing the external argument from the verb. In *Phrase structure and the lexicon*, eds. Johann Rooryck and Laurie Zaring, 109–137. Dordrecht: Kluwer.

Kratzter, Angelika. 2000. Building statives, eds. Lisa Conathan, Jeff Good, Darya Kavitskaya, Alyssa Wulf, and Alan Yu, 385–399. Berkeley, Calif.: Berkeley Linguistics Society.

Kratzter, Angelika. 2014. Situations in natural language semantics. In *Stanford encyclopedia of philosophy*, ed. Edward N. Zalta, Vol. 2014 Edition.

Kripke, Saul. 1959. A completeness theorem in modal logic. *Journal of Symbolic Logic* 24: 1–14.

Kripke, Saul. 1963. Semantical analysis of modal logic. *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik* 9: 67–96.

Landman, Fred. 1992. The progressive. *Natural Language Semantics* 1 (1).

Landman, Fred. 2008. 1066. on the difference between the tense-perspective-aspect systems of English and Dutch. In *Theoretical and crosslinguistic approaches to the semantics of aspect*, ed. Susan Rothstein. John Benjamins.

Leech, Geoffrey. 1971. *Meaning and the english verb*. London: Longman.

Lewis, David. 1973. *Counterfactuals*. Blackwell Publishers.

Lewis, David K. 1986. *On the plurality of worlds*. Oxford: Blackwell.

Löbner, Sebastian. 2001. The present perfect in german: Outline of its semantic composition. *Natural Language and Linguistic Theory* 19 (2): 355–401.

Lundquist, Björn. 2008. Nominalizations and participles in swedish. PhD diss, University of Tromsø.

Marantz, Alec. 1997. No escape from syntax: Don't try morphological analysis in the privacy of your own lexicon. In *Proceedings of the 21st Annual Penn Linguistics Colloquium*, eds. Alexis Dimitriadis and Laura Siegel. *University of pennsylvania working papers in linguistics*, 201–225. Philadelphia: University of Pennsylvania.

McCoard, R. W. 1978. *The english perfect: Tense choice and pragmatic inferences*. Amsterdam: North-Holland Publishing.

McNamara, Paul. 2014. Deontic logics, Winter 2014 edn. In *Stanford encyclopedia of philosophy*, ed. Ed Zalta.

Mittwoch, Anita. 1988. Aspects of english aspect: On the interaction of perfect, progressive and durational phrases. *Linguistics and Philosophy* 11: 203–254.

Moens, Marc, and Mark Steedman. 1988. Temporal ontology and temporal reference. *Computational Linguistics* 14: 15–28.

Moltmann, Friederike. 2018. Natural language and its ontology. In *Metaphysics and cognitive science*, eds. A. Goldman and B. McLaughlin. Oxford University Press.

Myler, Neil. 2017. *Building and interpreting possession sentences*. MIT Press.

Narrog, Heike. 2012. *Modality, subjectivity and semantic change: A crosslinguistic perspective*. Oxford University Press.

Nauze, Fabrice. 2008. Modality in typological perspective. PhD diss, University of Amsterdam.

Owens, Robert E. 2001. *Language development: an introduction*, 5th edn. Boston: Allyn and Bacon.

Palmer, Frank. 1986. *Mood and modality*. Cambridge University Press.

Pancheva, Roumyana. 2003. The aspectual makeup of perfect participles and the interpretations of the perfect. In *Perfect explorations*, eds. Monika Rathert Artemis Alexiadou and Arnim von Stechow, 277–306. Mouton de Gruyter.

Pancheva, Roumyana, and Arnim von Stechow. 2004. On the present perfect puzzle. In *Proceedings of NELS 34*, eds. K. Moulton and M. Wolf.

Parsons, Terence. 1990. *Events in the semantics of English: A study in subatomic semantics*. Cambridge, MA.: MIT Press.

Perlmutter, David. 1971. *Deep and surface structure constraints in syntax*.

Pfaff, Alexander P. 2015a. Adjectival and genitival modification in definite noun phrases in icelandic. PhD diss, University of Tromsø.

Pfaff, Alexander P. 2015b. Genitival and adjectival modification in definite noun phrases in icelandic. PhD diss, University of Tromsø.

Pietroski, Paul. 2005. Meaning before truth. In *Contextualism in philosophy*, eds. G. Preyer and G. Peter, 253–299. Oxford University Press.

Portner, Paul. 1998. The progressive in modal semantics. *Language* 74: 760–787.

Portner, Paul. 2003. The temporal semantics and modal pragmatics of the perfect. *Linguistics and Philosophy* 26: 459–510.

Portner, Paul. 2009. *Modality. Oxford surveys in semantics and pragmatics*. Oxford University Press.

Poser, William J. 1992. Blocking of phrasal constructions by lexical items. In *Lexical matters*, eds. Ivan A. Sag and Anna Szabolcsi, 111–30. Stanford, CA: Center for Study of Lang. & Information.

Potts, Christopher. 2007. The dimensions of quotation. In *Direct compositionality*, eds. Chris Barker and Pauline Jacobson, 405–431. Oxford University Press.

Pylkkänen, Liina. 1999. Causation and external arguments. In *Papers from the UPenn/MIT roundtable on the lexicon*, eds. Liina Pylkkänen, Angeliek van Hout, and Heidi Harley, 161–183. Cambridge, Ma.: MITWPL 35.

Ramchand, Gillian. 2008. *Verb meaning and the lexicon*. Cambridge University Press.

Ramchand, Gillian. 2012. On the categories of v and p: Anatomy of a love affair. In *Stuttgart workshop on prefix verbs*. ms., University of Tromsø. ms., University of Tromsø.

Ramchand, Gillian. 2014a. Deriving variable linearization. *Natural Language & Linguistic Theory* 32 (1): 263–282.

Ramchand, Gillian. 2014b. Structural meaning and conceptual meaning in verb semantics. *Linguistic Analysis* 39: 211–247.

Ramchand, Gillian. 2016. Events and verbal decomposition. In *The oxford handbook of events*, ed. Rob Truswell. Oxford University Press.

Ramchand, Gillian. 2017. The event domain. In *The verbal domain*, eds. Irene Franco Roberta D'Alessandro and Ángel Gallego, 233–254. Oxford: OUP.

Ramchand, Gillian, and Peter Svenonius. 2014. Deriving the functional hierarchy. *Journal of Language Sciences. Language Sciences*, <http://dx.doi.org/10.1016/j.langsci.2014.06.013>.

Reichenbach, Hans. 1947. *Elements of symbolic logic*. New York: Macmillan.

Ritter, Elizabeth, and Martina Wiltschko. 2009. Varieties of infl: Tense, location and person. In *Alternatives to cartography*, eds. Jeroen Craenenbroeck Hans Broekhuis and Henk van Riemsdijk, 153–202. Mouton de Gruyter.

Ross, John Robert. 1969. Auxiliaries as main verbs. In *Studies in philosophical linguistics*, ed. W. Todd, Vol. 1, 77–102. Evanston: Great Expectations.

Sailor, Craig. 2012. Inflection at the Interface. ms., UCLA.

Schachter, Paul. 1983. Explaining auxiliary order. In *Linguistic categories: Auxiliaries and related puzzles*, eds. F. Heny and B. Richards, 145–204. Dordrecht: Reidel.

Schwarz, Bernhard, and Rajesh Bhatt. 2006. Light negation and polarity. In *Negation, tense and clausal architecture*, ed. Raffaella Zanuttini et al., 175–198. Georgetown University Press.

Simpson, Jane. 1983. Resultatives. In *Papers in lexical-functional grammar*, eds. Beth Levin, Malka Rappaport, and Annie Zaenen, 143–157. Bloomington: Indiana University Linguistics Club.

Smith, Carlota S. 1991. *The parameter of aspect*. Vol. 43 of *Studies in Linguistics and Philosophy*. Dordrecht: Kluwer.

Sportiche, Dominique. Division of labor between merge and move: Strict locality of selection and apparent reconstruction paradoxes. ms., UCLA.

Stowell, Tim. 2004. Tense and modals. In *The syntax of tense and aspect*, eds. Jacqueline Guérion and Jacqueline Lecarme, 621–636. Cambridge, MA: MIT Press.

Svenonius, Peter. 2004. On the edge. In *Peripheries: Syntactic edges and their effects*, eds. David Adger, Cécile de Cat, and George Tsoulas, 261–287. Dordrecht: Kluwer.

Svenonius, Peter. 2012. Spanning. ms., University of Tromsø, <http://ling.auf.net/lingBuzz/001501>.

Taylor, Barry. 1977. Tense and continuity. *Linguistics and Philosophy* 1.2: 199–220.

Varasdi, Károly. Worlds, events and inertia. to appear, *Journal of Logic, Language and Information*.

Vlach, Frank. 1981. The semantics of the progressive. In *Syntax and semantics, vol. 14: tense and aspect*, eds. Philip Tedeschi and Annie Zaenen, 271–292. New York: Academic Press.

von Fintel, Kai, and Anthony S. Gillies. 2010. Must...stay...strong! *Natural Language Semantics* 18 (4): 351–383.

Wasow, Thomas. 1977. Transformations and the lexicon. In *Formal syntax*, eds. Peter Culicover, Thomas Wasow, and Joan Bresnan, 327–360. New York: Academic Press.

Werner, Tom. 2006. Future and nonfuture modal sentences. *Natural Language Semantics* 14 (3): 235–255.

Williams, Edwin. 2003. *Representation theory*. Cambridge, Ma.: MIT Press.

Wiltschko, Martina. 2014. *The universal structure of categories: Towards a formal typology*. Cambridge, England: Cambridge University Press.

Zamparelli, Roberto. 2000. *Layers in the determiner phrase*. New York: Garland.

Zanuttini, Raffaella. 1992. Syntactic properties of sentential negation. a comparative study of romance languages. *Dissertation Abstracts International, A: The Humanities and Social Sciences* 52 (7): 2536.

Zanuttini, Raffaella. 1997. *Negation and clausal structure: A comparative study of Romance languages*. New York: Oxford University Press.

Zimmerman, Thomas Ede. 2000. Free choice disjunction and epistemic possibility. *Natural Language Semantics* 8 (4): 255–290.

Zucchi, Alessandro. 1999. Incomplete events, intensionality and imperfective aspect. *Natural Language Semantics* 7: 179–215.