Its a Process *and* an Event

INSTITUTE FOR LOGIC, LANGUAGE AND COMPUTATION

Its a Process *and* an Event

Perspectives in event semantics

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de
Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof.dr. J.W. Zwemmer
ten overstaan van een door het college voor
promoties ingestelde commissie, in het openbaar
te verdedigen in de Aula der Universiteit
op woensdag 28 februari 2007, te 14.00 uur

door

Darrin Louis Hindsill

geboren te Los Angeles, Verenigde Staten van Amerika.

Contents

Acknowledgments

The process of finishing a PhD thesis was expected to be difficult, but through a variety a factors, it has been not only that, but multiplied by ten. Indeed, at certain points I never thought that I would be writing these lines. That they are written owes a great deal of gratitude to my supervisor, Michiel van Lambalgen, who convinced me in the midst of illness that I did have it in me to complete the work. Of course, he also deserves the more ordinary thanks (in his ordinary role as supervisor), for guiding me through the procedure, slowly forming what became a manuscript before my scarcely believing eyes.

I am also grateful to the ILLC administration for making my (extended) stay possible, again, without which I wouldn't have been able to reach this moment. Though I have been unable to show my face much in the last two years, I would also I like to thank everybody on the second floor of the Philosophy department. Even if I could not be a regular attendant, the talks I attended and discussions I had with my colleagues were certainly fruitful. It is also most appreciated that fellow AIO's would keep me informed as to happenings, even if I was notable by my absence. Special thanks is owed to Fabrice Nauze for Samenvatting assistance.

A similar gratitude is owed to my friends and flat-mates for support and help any time I needed it. So this is also for Karin, Cecilia, Maarten, and Sanne. The same applies to Joes as well, not only for the former, but also for descending from the airy realm of virtual design to the corporeal world of print.

Finally, I would also like to thank my family and friends in Los Angeles, especially my parents, who never wavered in their confidence, even if I did.

It is likely that I have left one or more people out. If so, and you are reading this, please accept that this is only my ignorance and faulty memory.

Chapter 1

Preamble

1.1   What do we really 'know'?

It is perhaps best to begin with a simple, but surprising, example that serves to illustrate a main theme of this thesis. It is generally held that certain types of statives resist the English progressive. The most typical of theses statives is *know*. A browse through the literature and a query of a the class of native English speakers involved in semantics or linguistics would certainly hold the following:

(1)   *I am knowing the answer.

The * should be read here as not merely semantically anomalous, but downright ungrammatical. But why should this be so? Certainly, the progressive construction is morpho-syntactically grammatical, so the deciding factor here is that something like knowledge in the philosophical sense cannot be seen as an activity,[1] but rather the state of an individual. Moreover, in the above example,[2] the object of knowledge is most likely a universal truth, and there is no context given, other than an 'out of the blue' statement of a piece of someone's knowledge.

This view in the semantics literature goes back at least to (Vendler, 1967) in his classification of English verbs into four discrete aspectual classes. Since that time, while many so-called stative verbs have been shown to appear quite commonly in the progressive (such as *sit, live, resemble*, etc.), the view that *know* cannot appear in the progressive has held fast. Nevertheless, there are some attested uses extant in the literature. (Huddleston, 2002b) states that progressive *know* "is just about restricted to waxing/waning case (*He claims that fewer and fewer students are knowing how to write English when they come up to university.*)" (p. 170). (Baker, 1995) has a waxing version:

---

[1] For now the reader can rely on their intuition for the meaning of states and activities. They will be defined in the next chapter.

[2] Which should certainly be said to be infelicitous or very anomalous to say the least.

(2)   Dana is knowing more and more of the answers as the course progresses.

This is about as far the acknowledgement of exceptions goes, but in neither example is it a progressive where someone knows a particular fact. In the first case it is a description of a series of students capacities, and the second the growth of knowledge about a specific subject. Keeping matters intuitive, both of the examples present something dynamic and activity-like, making them a prime candidate for the progressive. But (1) is knowledge of a single fact, and while one can, unfortunately, forget what one knows,[3] the knowledge of a fact during a particular moment in time hardly seems dynamic. Perhaps, then, this restriction on progressive *know* is correct after all.

But what if in discourse, the notion of knowledge plays a slightly different role? That is, aside from quizzes and nerdish braggadocio, people very rarely list the facts they know. Rather, if they make a point of knowledge of a particular fact, it is usually for a reason relevant to the current situation. This last point is the beginning of an insight that can then make predictions about more contingent variants of sentences such as (1).

(Goldsmith & Woisetschlaeger, 1982) modifies the notion of state as something that can be viewed in either a structural or phenomenal way. Some states can be seen as an inherent part of what they are predicated of, while some happen to be merely temporary states that are subject to change. Their famous example is that of the verb *stand*, and how the speaker construes it. Imagine that the statue's placement in both sentences has the same spatio-temporal duration.

(3)   a. The statue of Tom Paine is standing at the corner of Kirkland and
        College (but everybody expects the new Administration to move it).
        b. The statue of Tom Paine stands at the corner of Kirkland and College
        (and nobody thinks the deadlocked City Council will ever find a proper
        place for it.)

In the first case the statue's position is seen as temporary and the use of the progressive is allowed. In (3-b), the statue is, apparently, never going to move again, and its position is construed as more permanent, making the use of the simple form of *stand* much more felicitous.[4]

Back to the matter at hand, could we not then predict that when the import of the knowledge is relevant to a specific situation that is temporary, the progressive of *know* could be used? The answer, contrary to many native English speakers of a certain social class and background[5] is a resounding 'Yes'. Conventional biases aside, this isn't exactly the case. The following is taken from an interview with

---

[3]Again bringing a 'dynamics' of knowledge up.

[4](Comrie, 1976a) contrasts *The Sphinx stands by the Nile* with *Mr. Smith is standing by the Nile*. Presumably a tourist on a package holiday, here it is quite obligatory to use the progressive form of *stand* (aside from a narrative context).

[5]Including my own.

a comic book writer who wants to turn his creation into a screenplay and also direct the movie. This is his recollection going into a meeting with the producers. There is a plan to turn the comic book into a movie, but his role is still uncertain.

(4)   I said, 'let me write it and I'll prove that I'm the guy who can direct it.' Meanwhile, *I'm knowing* they're probably going to bump me and go get Ridley Scott.[6]

In this context, the use of progressive *know* doesn't sound ungrammatical at all. Importantly, it is the impact of the knowledge of a certain fact[7] that will have little relevance once the producers decide on personnel. In a sense, it can be said that his 'knowledge' is temporary and therefore there is no reason not to use the progressive (aside from conventional proscriptions, of course).

Stative verbs are also used in the progressive to express a degree change in the state or a series of entries into the state, or construing the state in an 'activity' perspective.[8] This will remain unexplained for now, but the reader is invited to make their own judgement on the acceptability of the following examples. These examples are much more similar in spirit to the examples of (Huddleston, 2002b) and (Baker, 1995). Importantly, however, they are actual examples by native English speakers who most likely could not care less as to whether cognitive stative verbs can be used in the progressive.[9]

(5)   a. "How could you not get on a roll with the way we started?" The confidence was up, and when I was taking the shots, *I was knowing* they were going in, not hoping they were going in.[10]
b. In the face of all that, I knew I could no longer embrace what *I was knowing* more and more to be a fiction.
c. One moment I could concentrate on any topic, and the information was coming to me, through my whole being. *I was knowing* things directly and at the same time I was very aware of everything around. I thought, this is silent knowledge.

For me these are all a bit less felicitous than the first example, but as descriptions of changes in one's knowledge, whether through a mystical experience, the sporting equivalent or confrontation with new facts that leads to a gradual reassessment, they fit in with the structural/phenomenal distinction and the application of the progressive.

---

[6]From a 13 Feb. 2003 interview with director/writer Mark Steven Johnson taken from joblo.com.

[7]The producers plan to hire another, much more famous director.

[8]See Chapter 2 for an explanation of these.

[9]These are message board postings and transcripts of interviews.

[10]This is a quote taken from an interview with a basketball player who was 'in the zone', the state of altered consciousness that athletes sometimes have, where everything they do is successful.

In fact, in some dialects and uses, progressive *know* has become something rather standard. A cursory glance at Hip-Hop sites (including song lyrics and messageboards) shows that *am knowing* is fairly standard and not confined to such specialised contexts as seen in the above. While this is not meant to be a scientific survey of the usage in various dialects, my impression is that the 'discourse-relevance' use of progressive know is quite common in certain varieties of African American English, Southern American English (a close relation of African American English) and Indian English. Moreover, it is certainly not a new phenomenon. In a novel of colonial Australia,[11] a character utters

(6)   I just cudn't be going, that do be all, me jewels. I was knowing it in me heart as soon as I lift. Poor Alec hadn't a word to throw to a dog.[12]

Certainly, it is most likely standard English that will rule such uses out.

1.1.1   What can we know from *know*?

There is a lesson to be learned from this. In semantics, the practise of putting stars before supposedly ungrammatical sentences needs to be given less prominence. Instead, the viewpoint should be reversed. The semanticist should formulate the theory and constraints in such a way that there are still predictions, but predictions of a different kind. If an example of a construction is indeed felicitous,[13] it might be better to then have the theory predict what kind of meaning it should have, rather than hastily predict them ungrammatical. In the case of *know*, using (Goldsmith & Woisetschlaeger, 1982)'s insight, it is a simple matter to then wonder that if *know* could be used in the progressive then it will have properties that are generally associated with using a stative verb in the progressive. One then needs to imagine cases where knowledge can be temporary or changing, and the felicitous examples follow.

Of course, this should not be taken to mean that anything goes. For example, any random string of phonemes will not be a sentence of English, and even a random string of conventional English words do not a sentence make. The point is rather that where a semantic distinction is at issue, it is better to think not in terms of grammaticality but in terms of context and possibilities of interpretation. The notion of the 'star' came out of theories of grammar that attempted to describe a syntax that[14] would only produce grammatical sentences, with 'stars' indicating sentences that the grammar would not produce. With semantics involved, an ungrammatical sentence could merely be a sentence out of context

---

[11]*Mistress Pat: A Novel of the Silver Bush* by Lucy Maud, 1935.

[12]This may need translating – *I just couldn't go. That's all it is, me jewels* (term of endearment). *I knew it in my heart as soon as I left. Poor Alec couldn't talk me out of it.*

[13]At least used by some sub-group a language's speakers.

[14]Performance issues aside.

that, once imagined in alternate construal, is perfectly felicitous and an example of a construction that is often used.

1.2 The Plan

In this new light, it is now time to discuss the body of the thesis. The semantics focussed upon is confined to that of events and their reporting. This includes the notion of eventuality types and their expression in tense and aspect. The next chapter is indeed such an introduction, and looks cross-linguistically at the various ways events are reported and construed, as well as making comparisons and connections between the sundry languages in a search for what can be seen as a ground for the disparate phenomena.

As already briefly demonstrated in the previous section, what emerges is a tension between a default way of lexicalising certain situations and concepts and what happens when a perspective shift or unorthodox construal has an impact on interpretation. It turns out that the interpretative process can be seen not as random and anarchic, but the effect of a particular construction (tense/aspect or adverbial) that shifts the meaning of event-type in a patterned and predictable way.

Having established eventuality types and the both the similarities and differences across languages in their construal and reporting, the question then arises as to what lies behind this. This question is looked at from the domain of psycholinguistics, both in the way different kind of events are constructed psychologically. Then we see how children learn to both construct event-types, notions of time, as well as eventually learn how to properly use the tense and aspect mechanisms in their languages. While no grand conclusions are drawn, combining the following two chapters together builds a picture of how humans psychologically view events and tense, how it develops, and a partial picture of how this appears in (a small subset) of the world's languages.

These two chapters also have a more (benignly) sinister motive, on top of general elucidation. They are also the prelude to the more formal chapters that use the formalism of the Event Calculus to formalise nominal, causal and perception constructions. More broadly, this typology of events and their psychological grounding is to show that this specific formalism (Lambalgen & Hamm, 2005), is motivated to take these notions seriously. Both the construction of events as well as tense and aspect constructions use formal concepts and mechanisms that can be seen as analogues as to what goes on psychologically, that is the building blocks of what is needed to conceive of, and talk about events in all of their temporal guises. Secondly, it takes the notion of construal and perspective seriously, and thus rather than a theory that sets out to limit the use of language, it allows predictions in a principled way, to show what happens when a given event-type

is construed differently.[15]

The latter three chapters focus on specific semantic topics. Two of them are about specific construction families in English of nominalisation and perception-verb complements, while the third a more universal exploration of the notion of cause[16] in the worlds languages. A basic goal of this is to extend the coverage of the event calculus to complements and various types of causal constructions. The chapter on nominalisation is a reworking (and much shorter version) of (Hindsill, 2001) with a more nuanced (and somewhat surprising) look at various nominal distinctions in English. All three have in common a look at *ing* in its various manifestations in English and its impact on tense and aspect.

While it could be considered a truism that one receives a certain delight in (eventually) coming up with a proper formalisation, it has more important uses. A closer look at the English perception verbs *see* and watch sheds light on the notion of event as well as demonstrates some rather fascinating interactions between matrix, complement and aspect. The chapter on causatives and resultatives shows how using the EC formalism, along with the interpretative principle sketched above, predicts how certain sentences, typically considered in the literature as ungrammatical[17] can have an interpretation renders them perfectly felicitous. More importantly, there are attested examples of exactly these interpretations. The constraint is then on the family of causative and resultative constructions that are formalised. One can then take the presumed 'ungrammatical' sentences, slot them in the construction and see what interpretation comes out.

The Interlude section, a semi-formal introduction to the Event Calculus demonstrates examples of how this 'slotting-in' can be accomplished in a systematic way, via a inferential process that unifies 'anomalous' lexical material with parts of the event structure. Thus, the meaning of coerced utterances can be predicted in a principled way.

1.2.1   How to read

The intended audience of this thesis is for those with a philosophy of language and semantics background. It is for this audience that the following 'guide to reading' is given. Chapters 2 and 3 are primarily foundational. The former is a discussion of event-types, their interaction with aspectual constructions as well as coercion and event structure. The topic of the first section – Aktionsart – will mostly be familiar to the reader. But as it discusses the topic with a wide variety of languages, it may still be of interest to those who are only familiar with the concepts as they apply to English or other European languages. The second section is an exploration to the many exceptions to the defaults given in the first

---

[15]Later to be called 'coercion'.

[16]Though again, certain English-specific causative and resultative constructions are given a special examination.

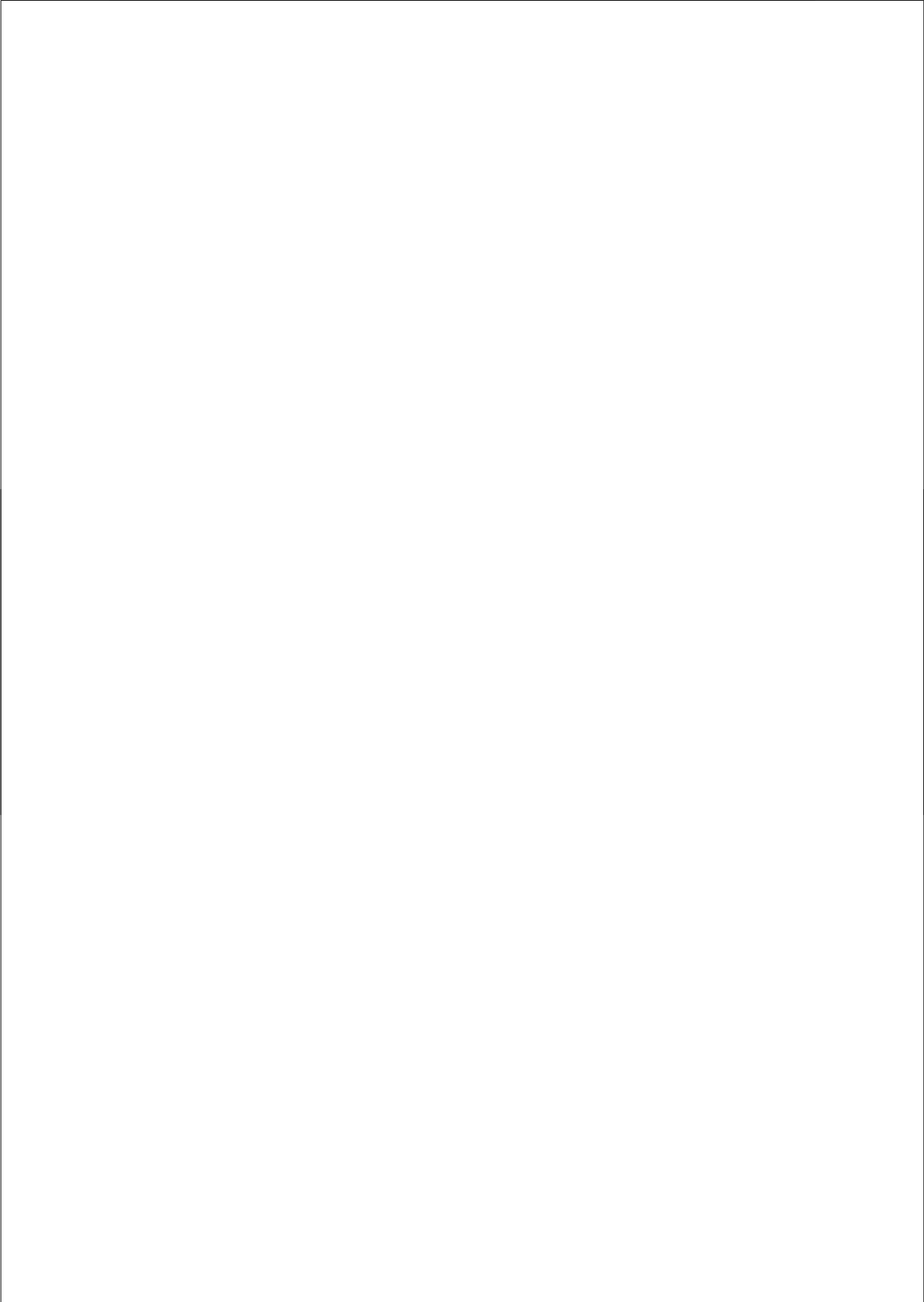[17]These predictions are for English constructions only.

section, as well as the first foray into the semantic structure of events. Again, this may be redundant for some readers, but for those for whom it is not, it provides a critical background of the later chapters, both theoretically and formally.

Chapter 3 is a exploration of the psychological parallels of the material in the chapter previous to it. It is important to note that this is background and elucidation for those of a semantics background, and in no way intended to be original work in the field of psycholinguistics. Thus the reader should not expect the methods or organisational principles of this field to be used. The first four sections are an examination of the psychological grounding of event structure and play a significant role in the three chapters that follow, especially in the way events are interpreted and expressed. The final sections of the chapter look at differing views on the acquisition of tense and aspect. While this and the sections on the psychology of event structure can be seen as a grounding for the formalism, the acquisition sections can remain unread as they have but a small impact on the comprehension of the later chapters. However, aside from being of interest to the reader who knows little about the topic, it has its own payoff in serving as background to a discussion of (Wagner & Carey, 2003) (Section 7 of the chapter). These experiments go beyond the topic of tense and aspect acquisition, and return to the question of event interpretation from a rather interesting angle.

Following this comes the Interlude, which (along with the Appendix) provides the reader with an introduction to the Event Calculus, with the Interlude meant to be understandable to those unfamiliar with the formalism. The hope is that, even if the gritty details of the formal sections of the later chapters cannot be grasped, the reader will now have enough of a background to follow the general outline of them.

The next two chapters – 4 and 5 – can be considered the core of the thesis. Both begin with a good deal of background material, building up what is necessary for the various formalisations. Note that observations due to other authors are signalled by references, and that original contributions, even if not explicitly stated, are to be found where references stop. While the final sections of both chapters contain the bulk of the contributions, scattered throughout the earlier sections is commentary and critique of other authors' work, both from a theoretical and empirical perspective. Finally, while Chapter 5 does explore a good deal of linguistic typology in order to understand the linguistic notion of a causative construction, the same caveat applied to Chapter 3 applies.

Chapter 6 on English *ing* nominalisation is a slightly looser fit, but has connections with Chapter 4 in its examination of the varied uses of *ing* in English, as well as providing yet more insight into perspectives on events. Here, the major contributions are primarily empirical questioning of standard notions, as well as speculation as how these empirical phenomena can fit into the notion of event semantics advocated. The brief formal sections at the end of the chapter provide hope that indeed this can be done.

Chapter 2

Lexical Aspect

2.1 Event Typology

2.1.1 Introduction

The language, Kalam, spoken in the Highlands of Papua New Guinea, has but ninety verbs, only thirty of which are used with any regularity. According to (Pawley, 1987) these are verbs in the usual sense in the term, in that only they may morphologically mark tense (both relative and absolute), aspect, mood, number, etc. (via suffixes).[1]

But despite the apparent paucity of verbs, a Kalam speaker is able to express virtually any action or situation compositionally by combining verbs in serial constructions (see footnote) or in combination with the wider class of nouns. For example, the verb *nŋ-* can express sensing, seeing, hearing, thinking, remembering, being intelligent, etc. Though Pawley does not translate this verb with one, more general gloss, it seems reasonable to make the connection. (Frawley, 1992) notices that this particular verb is closely related to the notion of an internal state.[2] A particular internal state is expressed as in the following examples:

(1)   wdn nŋ-   *see*
      eye   Internal state

(2)   gos   nŋ-   *think*
      mind Internal state

---

[1]Verbs can occur in serial verb constructions, with all but the final verb occurring in the bare form, and the final verb receiving inflection. Interestingly, the final verb can also mark relative or absolute reference of subject, i.e. whether or not the subject of the final verb is the same as the non-final. Also note that these serial verb constructions are to be regarded as strings of separate verbs, rather than idiomatic constructions.

[2]The notion of 'internal state' and state more broadly conceived will be elaborated shortly.

(3)   wsn   nŋ-   *dream*
       sleep Internal state

(4)   d   nŋ-   *feel (by touching)*
       touch Internal state

Note that (1) - (3) are nouns combined with the internal state verb (or as Pawley translates it, 'Perceive'), while (4) is an example of a serial verb construction, with another 'general' verb, translated in this context as 'touch', combined with the internal state verb to derive *feel*.

   Nor is ŋ- the only Kalam verb with such a general meaning. The verb *d-* (seen translated as 'touch' in (4)), has translations by Pawley for such verbs as *get, hold, handle, grasp, capture, possess, finish*. For example, when combined with the verbs for *go* and *come*, we obtain the serial verb construction for the English verb *fetch*.[3]

   As with ŋ-, Pawley would like to give a general gloss to *d-* that can didactically be used to make a point about the close relationship between verbs and events. In this case *d-* is a verb denoting 'resulting events'.[4] And indeed, the various translations given to *d-* do involve results of an action (e.g. *get, take, bring, finish*). That *d-* is also translated as *hold* or *possess*, may be due to the fact that the result of the above sorts of actions often involves possession or holding. The next example serves to illustrate this:

(5)   B   tw   dp
       man   axe he-took
       The man took hold of an axe.

For actions themselves, a different verb is used – *g-*. Pawley translates this verb with such glosses as *do, act, make work, create, function* and *cause*. For example, the English VP *build a house* would be translated as:

(6)   kotp   g-
       house create

while 'feeling hungry' also uses the same verb.

(7)   yp   ywan   gp
       me   hunger it-acts
       I feel hungry

---

[3]As another nice example of Kalam compositionality, if the *come*(back) component of the construction is removed, the result is *d am* (*get go*), which is translated as 'take', while *d ap* (*get come*) is translated as *bring*.

[4]As this introduction is used only to set up a conceptual event typology later in the chapter, any terminology used here is for didactic purposes only.

(Frawley, 1992) simplifies matters a bit by claiming Kalam verbs denote only "states, processes, or actions", denoted by *ng-, g-* and *d-* respectively. His aim is to find a limiting case for the claim that verbs are semantically motivated by events. This corresponds with the pre-theoretical notion that verbs are 'action' words and therefore denote some kind of event. For example, the English verb *build* can be seen as an event which involves some sort of process over time that, if successful, creates an object. In this case, there are some 'event' notions of a process, and something changing (an object coming into existence) along with the specific lexical material and concepts associated with building. But what Frawley means by a 'limiting case' is that the above Kalam verbs not only denote events, but denote undecomposable events. That is *g* denotes only action, and needs to be incorporated with some other verb or noun to give it lexical content. The claim is that what has been found is the Holy Grail of primitive, universal semantic event types existing in a remote New Guinea language.

Indeed this claim is rather robust, and can be seen in the formal semantic literature in the use of a Davidsonian Event Variable *e* as an additional argument of verbal predicates.[5] With Kalam, there is a rather discrete division of the 'event space' that gives actual content to the oft used term *verbal aspect*. Indeed, the point of bringing up this particular language is that the conceptual event distinctions made by typologists seem directly encoded in Kalam syntax.

My reading of Pawley's paper leads me to think that the claim that these three verbs perfectly divide the event space is too strong. Note that there are also 'general' verbs to cover transfer of possession, location, stabilization and destabilization as well as specific verbs such as *come, cut, bake, eat*, etc. Nevertheless, there is an uncanny resemblance between Kalam verbs and standard event typology, as will be seen in the following sections.

But first, there is one more example from Kalam that bears looking at. In Kalam, should there be a beneficiary of a certain action, it cannot be encoded in the same clause as the action. For example, in English, one could use a *for* clause to encode the beneficiary of an action, such as in the sentence, *I am building a house for you*. Kalam requires at least two clauses to express the same relation – one for the action (the creation of an object, or bringing it, etc.), and one for the actual transfer of possession. Furthermore (and more interestingly for our purposes), in a situation such as a house built for someone else, three clauses are needed:[6]

---

[5]This idea originating in (Davidson, 1967). The Davidsonian event variable does not, however, distinguish between different types of events. Alternatively, one could argue that time is an inherent part of the verb and derive events from this – this is much more preferable in the author's view. This is not to deny the close connection, of course, between verbs and events.

[6]In the gloss below, *SS* merely means that the two clauses have the same subject. In this case the person who has built the house is the same as the person who intends to give it to the beneficiary.

(8)   kotp gy,   np ñnp   gspyn
        house having-built-SS you intending-to-give-SS I-am-doing
        I am building a house for you

Aside from the separate clause for the beneficiary, there are two separate clauses
to encode the speaker's building of the house. There is both the (soon to be)
resulting state of having built the house, along with the current action. This is
also the case for instrumental clauses. For example, *The man split the wood with
an axe* (which can also be split up in English, e.g. *The man used an axe to split
the wood*) would be translated as

(9)   b   tw   dyt,   mon tb lak-p
        man axe having-taken-SS wood cut he-split

Toward the end of this chapter, it will be shown that this is strikingly similar to
a standard event structure and typology.

   As noted briefly above, there is a standard notion that verbs are inherently
tied to *events* in some way, and that different verbs may denote different types
of events. This idea seems to have a high correlation in Kalam, where the kind
of event denoted can be read off from the serial verb construction. For example,
as we saw above, any verbal construction with *ŋ-* will be some sort of 'internal
state'. More generally, it is quite common to see the literature full of reference
to 'stative verbs', 'activity verbs', etc.[7] This is what is known as *Lexical Aspect*
or *Aktionsart*. In the next section we will look at the standard event typology, as
well as what an event (more generally) consists in. However, it will also be seen
that a general notion of a strict mapping from verbs (or VP's) to event types is
not tenable, certainly for languages such as English. Instead, event types should
be seen as a conceptual typology that may be represented in language rather
closely (as in Kalam) or less so (as in English).[8]

2.1.2   Aktionsart

In the last section it was claimed that verbs can be directly associated with
different sorts of situations. That is, certain verbs may describe situations where a
subject has a certain property (e.g. *She is beautiful*), or describe situations where
an action take places (e.g. *John is running* or *John is building a house*), or even a
situation that describes an 'instant' transition (e.g. *The window shattered*). Each
one of these situations can be seen as *events*, and moreover as being instances of
certain *types* of events. It is the various types of events that will be looked at in

---

[7]For example, *think* would be considered 'stative', while *run* would be an 'activity'.

[8](Freed, 1979) notes that for English, the event-type can not be read off VP's *per se*, but is
a property of the whole sentence.

this section. As previously mentioned, this is often called *Aktionsart*[9] or *verbal aspect*.

In contrast to grammatical aspect, the typology in *Aktionsart* should be seen as a non-linguistic conceptual typology. Though it is often called *lexical* or *verbal* aspect, the categories cannot be systematically mapped to lexical items in a language such as English.[10] Instead, I tend to see a language such as Kalam, not as evidence for *verbal aspect per se*, but instead as linguistic evidence for a non-specifically linguistic conceptual system.

This conceptual system can be traced as far back as Aristotle (e.g. categories in the *Metaphysics* and *Nicomachean Ethics* of *enérgeiai*, *prâksis* (*doing*) and *poíēsis* (*making*)).[11] These conceptual distinctions were then taken up in the 20th Century (in a more linguistic vein) starting with Ryle, Vendler, and Kenny.[12] In what follows I will give the basic Aktionsart distinctions with some modifications, discuss their 'default' representation in language, as well as a rather basic idea of event structure taken from the work of (Moens & Steedman, 1988), and the similar ideas of Freed. It should be noted that in defining and contrasting the basic event types, various linguistic 'tests' will be used. These are meant to be diagnostics and thus cannot be seen to be without exception, and are intended mainly to highlight the conceptual distinctions. Moreover, the fallibility of the tests does serve to expose the default mapping from lexical item to event type, and thus even the exceptions are illustrative. With that in mind, we will now look at some basic event distinctions starting with state vs. non-state.

Stative vs. Non-Stative

It may seem rather strange at first (in common sense terms) to include the notion of a 'state' in a broader notion of event. Certainly, we typically think of events as ranging from World War II, a wedding, or even a sneeze.[13] The common notion behind these things is that something actually happens – which is exactly what is not the case with a 'state'. Nevertheless, the notion of state is integral to a

---

[9]Confusingly, *aktionsart* is also used in a wider sense encompassing much of 'grammatical' aspect including resultative, ingressive, continuative aspect. We leave the examination of these to another section.

[10]Though arguably most VP's of English can be said to have a 'default' event-type meaning – more on this later.

[11](Binnick, 1991) also notes that there seems to be much Indian work on aspect as far back as the 5th century B.C.E. However, this probably had less influence on western linguists and philosophers than did Aristotle.

[12](Vendler, 1967), (Ryle, 1949), (Kenny, 1963).

[13]It is for this reason that the term 'eventualities' is often used as a technical term. In common parlance there is a notion that certain types of eventualties to be called 'events' must also happen rather compactly. There is a fairly common saying – 'This is a process, not an event' meaning that the desired outcome will take some time rather than a day or two. Politicians will often use such an expression when their beloved programs are not yet having the desired effect.

more 'extended' notion of event. Furthermore, the same situation ontologically may be viewed as 'stative' or 'eventive' (in the sense of something happening), and therefore it is imperative to include states in a typology of events.

States are typically denoted by so-called 'stative verbs' such as *love, live, be X, resemble and know.* They are conceived as a homogeneous and durative situation. Moreover, the internal temporal nature of a state is undifferentiated. That is, all stages of a state can be said to be the same, and have a certain inertia about them. A now rather standard description comes from (Comrie, 1976a) (p. 49):

> With a state, unless something happens to change that state, then the state will continue...With a dynamic situation, on the other hand, the situation will only continue if it is continually subject to a new input of energy...To remain in a state requires no effort, whereas to remain in a dynamic situation does require effort.

At this point, and in everything that follows, it needs to be emphasised that any claims made about states or any other sorts of events are not to be taken as claims about reality, but rather our way of organising reality. For example, a prototypical state might be *Mary is beautiful* or *Mary resembles her mother.* If we apply reality to some of the criteria for states listed above, then these are not states at all. For example, appearance does change slowly over time, and one might want to make the rather sexist observation that being beautiful takes a lot work (i.e. input of energy). Furthermore, it is arguable that there are no actual spatial or temporal boundaries in nature (aside from cognitive perspective). That is, is there always to be a certain point where Mary is beautiful and another one where she stops being so? Probably not, but as we shall see in this section as well as those on coercion and aspectualizers, linguistic evidence does show that we can conceive of the world being this way.

That said, we can now go on to a more detailed description of states. As mentioned above, they are conceived of as unchanging over time as well as extended in time, and persistent (i.e. do not need to be actively continued). One test for them is that when used (in English) in the present tense, they tend to be incongruous with the *progressive* aspect. Thus:

(10)   a. John is tall.
       b. Mary knows the answer.
       c. #John is being tall.
       d. # Mary is knowing the answer.[14]

The '#' before (10-c) and (d) is not an indication of their ungrammaticality *per se*, but rather their infelicity in the an ordinary context, where the copula and *know* are typically stative. Certainly (10-c) is imaginable in a situation where

---

[14]But see the introduction to this thesis.

John is standing on tip-toe or wearing shoes with extremely high heels; however, a vital statehood criterion would then be lost– the 'tallness' would no longer be seen as inherent and persistent. Instead he would be *doing* something to make himself tall, and continue in this action in order to keep himself tall.

Sentence (10-d) is a more interesting case. *Know*[15] is considered almost a paradigm stative verb, and seen to be more resistant than most to progressivization. Nevertheless, I have found many examples that indicate that it occurs in both Indian English and some Southern and African American dialects. It is usually used to indicate the importance of the particular piece of knowledge at the time (e.g. *I was knowing at the time that X, but ...*). This is most likely due to general expansion of progressive marking across many English dialects.

Indeed this sort of extended use of the progressive has been described by (Goldsmith & Woisetschlaeger, 1982) as capturing a structural/phenomenal distinction grammatically marked in English by simple versus progressive marking. This is a kind of metaphysical perspective that treats a situation as either an inherent, permanent part of the world, or a transitory phenomenon (or at least the property's holding is important for a particular time rather than holding more generally). As an illustration, compare the two sentences:

(11)   a. This engine is running well.
         b. This engine runs well.

(11-a) is compatible with the engine having been ill-tempered for a while and only recently tuned up, whilst (11-b) generally means that it is an inherent property of the engine that it runs well.[16] However, the implicature in (11-a) that the engine had not been running well is easily cancelable. Imagine a situation where someone is trying to sell a used car.

(12)   How's she running? Oh, the car's running perfectly – always has done.

In this case, there is no change in the engine's (or car's) mechanical properties, but rather the speaker wishes to emphasise that this property especially holds at the moment. This is most likely what is going on with the progressive use of *knowing*, and as Comrie notes, for English, many stative verbs such as *live* and *stand* that may not occur in most languages with progressive marking (for languages that have such a grammatical distinction), occur regularly in English with progressive marking, when these are 'temporary' states.

A little speculation would see the extension of the progressive as developing from the stative/dynamic distinction. Unlike Spanish, where the progressive is

---

[15]This is *know* in the *weten* sense. As *acquaintance*, using progressive marking is perfectly ordinary. *I'm knowing you less and less the longer we are married.*

[16]In general, English simple tense will have this sort of meaning, along with generics (*Cows eat grass*) and habituals(*John smokes*), all of which are 'structural' perspectives. Compare with *Cows are eating grass, John is smoking – it's a beautiful day!*.

not obligatory for dynamic situations (*Juan cantá* can mean *John is singing*),[17] it
is for English. Moreover, stative verbs have their default expression in the simple
forms. Now, the conceptual distinctions between stative and dynamic invite the
inference to states (as persistent and often involuntary properties that are features
of the world) being permanent, while a dynamic situation that is not persistent
and needs constant energy to keep it going as being temporary. It then seems
possible that the grammatical marking itself then is picked up this distinction.

Of course not all states are even conceived of as permanent – along with
the temporary sitting and standing examples, there are states such as being sick,
which hopefully are quite temporary. Furthermore, in these cases one can only re-
ally say *I am sick* and not **I am being sick*. But, these can always be paraphrased
with a dynamic progressive periphrasis such as *I am feeling ill* or something sim-
ilar. In a language such as Spanish with two forms of *be* that roughly delineate a
temporary/permanent distinction, the form of copula greatly affects the meaning.
There is both *Soy enfermo* (I am a sickly person) and *Estoy enfermo* (I am feeling
ill). Using data similar to the above along with data regarding social roles and
temporary physical states (*She is the mayor of San Rafael*, *The clothes are dry
now*), (Croft, 2000b) modifies Goldsmith and Woisetschlaeger's construal of the
simple present/progressive dichotomy to capture only the inherent/generic and
habitual senses of the simple present.

Excursion – activities vs. states

Some languages will go so far as to morphologically mark the difference between
temporary and inherent states. Kobon[18] is the closest relative to Kalam and
shares about half of its vocabulary with it. The two languages, however, are
mutually unintelligible. For the purposes of this section, the way the language
handles 'stative' predication is at issue.

For standard adjectival complementation the adjective can either be incorpo-
rated into the noun or verb phrase or a clitic (*-bö*) which is copular) or a particle
rö (meaning *like*). This can be seen in the following examples (Davies, 1981)(p.
42-43).

(13)   Nipe bɨ   kub
         3s   man big
         He is a big man

(14)   Nipe kub rö
         3s   big like
         He is quite big.

---

[17]The progressive version would be *Juan está cantando.*
[18](Davies, 1981).

(15)   *Nipe kub
        *3s   big
        *He big.

(16)   Kumi kabö al̃i g-ö*p
        cloud stone red do-perf3s
        The sky is red.

(17)   Kumi kabö al̃i-bö
        cloud stone red-ADJR
        The sky is red.

(18)   *Kumi kabö al̃i
        *cloud stone red
        *The sky red

These are all fairly standard ways for languages to express predication, but what
is interesting is what happens when present-perfect marking is used with nominal
predication. Unlike adjectival predication, there is no required copular element.
Thus 'He is a man' is merely *Nipe bɨ* (he man). This is analogous to sentence
(14), where the adjective is incorporated into the noun, thus making a new noun
phrase.

  But, there is a verbal copula *mɨd* that is then inflected with the perfect tense-
aspect marking. If the predication is seen by the speaker as only of a temporary
nature then the present perfect copular coding may be used (but is optional), but
if it is seen as inherent, then it is not allowed.[19] [20] So (p.42):

(19)   Nipe kaunsol   (mɨd-öp)
        3s   councillor be-perf3s
        He is the councillor

(20)   Nipe bɨ   majö   (*mɨd-öp)
        3s   man mature *be-perf3s
        He is a mature man

---

[19]All examples of Kobon sentences are from (Davies, 1981) and the page numbers given refer
to this reference grammar.

[20]As statives are not allowed with present tense morphology (Davies, 1981) (p.171), present
perfect marking must be used if the sentence is to have tense at all, which explains the trans-
lations. Present tense seems only used for actions which, like Kalam use the verb *g*. Thus

(i)   Nipe wög g-ab
        3s   work do-pres3s

It becomes a bit more complicated if the situation is in the past. Suppose that in (20) the speaker believes that the addressee knows that the subject has died. If so, then (20) may be used (the present perfect marking is often used for simple past except in the case of durative but temporary situations (p.167)). But if the addressee is not aware of the death, then habitual marking must be added along with the copula:

(21)   Nipe bɨ   majö   mɨd-mɨd-öp
        3s   man mature be-habit-perf3s
        He was a mature man

At first glance, this patterning may seem a bit odd,[21] but the important thing to see is not that there is a grammatically realised distinction between temporary and inherent states (which is a common feature in many languages), but rather it is the way it is accomplished.

The use of present perfect marking to express a temporary state is conceptually quite reasonable. A state that is conceived as temporary, presumably begins at some point in time and is caused by an event. For example, in English, one could say for the same unfortunate situation

(22)   a. I am ill.
        b. I have got the flu.

where the first example is (hopefully) a temporary state, and the second a formulation in the present perfect.[22] In other words, as the state is construed as temporary, it is natural that the preceding, causal event would also be profiled. One could also construe the situation as an activity, and use the progressive to express, *I am feeling ill.*, which is exactly how it is done in Kobon (p. 179)

(23)   Yad nan   g-öp
        1s   thing do-perf3s
        I am ill

---

[21]Furthermore, I think it indicates that (Frawley, 1992) may have been a little to enthusiastic in trying to establish for the related Kalam that all verbs correspond to undecomposable primitives of Vendlerian event types. Certainly, stative predication is probably a little too complicated to warrant such a conclusion. Should Kalam stative making work more or less in a similar way, the 'stative primitive' verb only applies to a select number of internal states. More generally, it is possible that stative marking could be rather similar to what is described above. However, in favour of Frawley's claim, inherent stative marking in the relative, Kobon, involves no verbal usage (even copular), whatsoever.

[22](Lambalgen & Hamm, 2005) (p. 151) formalises the latter as a state that holds at speech time along with a call to the database to derive a causal event that initiated the state. See the Interlude following Chapter 3.

Weather states to work in the same way. Thus:

(24)   Igön g-öp
        cold do-perf3s
        It is cold.

Where 'cold' is a temporary state for the weather to be in, certainly in New Guinea. An interesting weather-related distinction between states and activities, cited by (Foley, 2005), occurs in the Maluku language of Taba. In Taba, there is the unaccusative *kabus* meaning wet and the unergative, *-mang* meaning dry.[23] As Foley notes (p.410) [24]

> Being wet is taken as the state an object is in, but being dry is the result of an unfolding change in state, i.e. drying out (remember that the language is spoke in a hot humid tropical area with high annual rainfalls!)

The above shows that there what one language grammatically defines to be a state, may be a process in other languages, often for the simple reason of cultural and environmental factors. But there are also deeper conceptual reasons in the demarcation between stative situations and dynamic situations – commonly referred to as processes and events.[25] We will look briefly at some more differences between stative and dynamic situations as well as certain pitfalls that come from trying to make too fine a distinction.

Conceptual differences between stative and dynamic 'events'

We have seen so far that some languages will make a lexical or morpho-syntactic distinction in treating temporary vs. inherent states, and also that some verbs can be marked as states in some languages and as activities (or the results of activities) in others. But the question then rises as to how far this process can go. Can any state be conceived of as an activity and vice versa?

First off, it is commonly assumed that a prototypical state is non-volitional, while the opposite is true for a dynamic situation. For example, one usually does

---

[23]The terms 'unnaccusative' and 'unergative' are not to be mistaken with their uses in the semantic literature to do with types of causatives. Here it means that the language has two sets of intransitive verbs that grammatically marks the subject as either an actor (unergative) or as undergoer (unaccusative). The former are taken to be action verbs and the latter, statives.

[24]Taba is spoken in East Timor and Indonesia.

[25]Here a bit of terminological confusion arises. On the one hand it is quite necessary to include all the Vendlerian types under the rubric 'Event', but there is also a specific type of event, that is also properly called an 'event' in contrast to a state or process. There certainly is wisdom in Bach's use of the term *eventuality* to cover all the Vendlerian types. I find this term rather cumbersome, and will be careful to make the two notions of 'event' distinguishable by context.

not choose to be tall, and nor is an individual free to choose to stop being tall. In contrast, if one is running, then this is normally by the volition of the subject, and they are free to stop running at any time. However, Binnick points out that there are a number of 'volitional' states, such as being a patriot or being a believer in God. In these cases, Binnick believes that 'volitional' states share a property with prototypical dynamic situations – that of being possibly 'gappy'. That is, it is conceivable that I could say I have been a patriot my whole life even if I have had a few dark moments of doubt. But it is rather strange to say I have been tall my whole life, if I mysterious shrank for a month. With a dynamic situation such as running, there is also a possibility of gappiness. Certainly it is true that I ran for two hours, even if twice I had to stop and tie my shoes.

Ultimately the conceptual distinction between states and dynamic situations is less clear-cut than it would seem at first. As noted in (Comrie, 1976a) (p.35), if the progressivizing test for stative situations is used to differ the static and dynamic, there must be awareness that not all languages with the progressive mark the same lexical items. For example in English and Spanish, one must put *rain* in the progressive (*It is raining, está lluviendo*), while the opposite is true for Icelandic. He also notes that while using *see* in the progressive is rather marked in English (*I'm seeing you there under the table*), it is perfectly acceptable in Portuguese. Again, the deciding factor is one of perspective. For a rather large range of situations, it is perfectly possible to view them as either stative or dynamic, and languages are free to choose how they classify them. Recall that in Kalam (ex. 8), hunger is a dynamic situation (using the verb *g-*), as it is hunger that is acting on the subject.[26] However, this flexibility is not open-ended; Comrie notes in a footnote that the whole classification (and certainly the progressive test) would fall apart if some languages classified such verbs *die*, *hit* and *kill* as default statives.

Comrie does attempt to get at the distinction via a notion of homogeneity of *phase* – that all phases of *John knows the answer* are identical to each other, whereas *John is running* will differ at various phases of the situation (for example having different feet on the ground, etc.) However, he then notes that a dynamic process like an oscilloscope emitting a pure tone at a certain frequency would be homogeneous in this sense. Thus, he reverts back to the quote above, and differentiates statives and dynamics based on energy input.

(Binnick, 1991) attempts a similar differentiation, by bringing in notions of both *stage* and *phase*. The claim is that states have no phasic structure. The difference between stage and phase (for dynamic events) can be seen as follows. In an accomplishment like *running a kilometre race*, two stages (which could be different parts of *running* in *running a race*) can be identical in just being *running*.[27] These stages can be gathered up into one phase that is the activity phase

---

[26]Of course, English can view this situation either as a stative or dynamic – *I am hungry* vs. *I am feeling hungry*, and it is perfectly natural to think of hunger growing over time.

[27]Generally, simple activities such as *running* are arbitrarily divisible. For example, take 10

of the accomplishment.[28] There is then a point at which the accomplishment is achieved, that is different in the sense that it represents a boundary 'beyond which the boundary qua accomplishment' cannot continue. For example, upon finishing a race, the runner can continue running for a bit (to cool off and also not let their muscles atrophy), but they are no longer running the race. While a state may have stages, there is no analogue to this notion of phase.

He also uses (Freed, 1979)'s aspectualising tests to distinguish activities from states. For example, one can stop being ill, but generally not finish being ill. But this sort of test also separates out telic from non-telic activities.[29] One can finish running a race, but generally not finish running, unless there was some telic endpoint presupposed of time or distance involved.[30]

So, I still don't see how this helps with distinguishing simple, non-telic activities from states. Comrie's examples of verbs that he couldn't imagine being classified as states are all telic (*die, kill*) or of the more instantaneous variety (*hit*). As he notes, *he's ill* cannot be taken as either *he's starting to be ill* or *he's about to be ill*, but *he's running (the race)* can be used even as the runner enters the stadium. Where it is only an atelic *running* activity, I am not sure which side the progressive takes. Certainly, *He's running in a minute* is fine as futurate, if someone is tying their shoes, and has yet to start running. But I would say that the *running the race* example is more of a true present progressive where the athlete has entered the stadium or just emerging from the starting blocks.

Regardless, this gets us no closer to conceptually dividing states and atelic activities. Both are uniform and both can be trivially gappy. This is a more general tendency for any atelic activity to allow gaps, but certain states allow this as well. Paraphrasing an example from Binnick, there is a difference in being in pain for a week and in *constant* pain for a week. It seems then, to use Binnick's terminology, both states and simple activities have no phasic structure (just stages), leading one to think that states and simple activities can be the most interchangeable of Aktionsart.

Thus far, we have seen linguistic and conceptual differences between stative and dynamic situations (and certain pitfalls in distinguishing them as well). Next we will move on to processes – dynamic situations that are extended in time, and then finally to what are often considered 'events' proper, achievements and points.

---

minutes of running. Any portion of that (ignoring rests from exhaustion) can also be running. But a random two minutes of running a kilometre is not *running a kilometre.*

[28]This will be explained in more detail in the Event Structure section. But for now, imagine that something like *run a mile* consists of both a *running* activity and the punctual event of reaching the mile boundary. The phase is then the entire *running* activity. The running activity itself can be construed as having stages.

[29]Telic ones are those with an inherent goal.

[30]Or as with states, if the *running* in question is a habitual activity that one has had enough of.

Processes: Activities and Accomplishments

Along with what may be seen as a primary conceptual opposition between static and dynamic situations, there is also a primary opposition within dynamic situations; this is commonly known as the difference between activities and accomplishments. As is the custom, we will first look at some linguistic data to illustrate the difference:

(25)   a. John ran a mile in under four minutes.
        b. Sheila wrote three letters in an hour.
        c. #John ran a mile for six minutes.
        d. #Sheila ate an apple for ten minutes.

(26)   a. John ran for twenty minutes.
        b. Sheila ate apples for two days straight.
        c. #John ran in twenty minutes.
        d. #Sheila ate apples in two days.

Note the contrast between the (a), (b) and (c), (d) sentences in (25) and (26). All of the above situations are dynamic (note that they all are felicitous in the progressive if they are in the simple past as well). But there is a crucial difference that is elicited by the use of the two different time adverbials – the notion of a goal. First coined by (Garey, 1957), this is the *telic/atelic* distinction (the *telic* situations being the ones that have an inherent goal or endpoint).

For example, merely *running* or *running in the park* has no inherent goal associated with it. Thus it is measurable (*for X amount of time*), but reaches no specific endpoint (i.e. one generally stops running but does not *finish* running). In contrast, with *run a mile* there is a predefined distance that was completed in a certain amount of time.

One can then differentiate two types of dynamic situations – most commonly called activities and accomplishments. Activities are the atelic, dynamic situations and include such things as *running, pushing a cart, reading, playing the piano*, etc. Accomplishments are dynamic, *telic* situations such as *building a house, crossing the street, flying to London, playing a sonata*, etc.

The idea of a goal can then show us why the temporal adverbial tests above can actually distinguish between the two event types. Of course, both *container* (*in x time*) and *durative* (*for x time*) adverbials measure out the time of the situation.[31] The difference is that two different type of situations are measured. With activities, the durative adverbial is used to measure the total time of said activity, whereas with a container adverbial (for accomplishments) the time measured is the time until the goal is reached. As activities have no inherent goal, an *in* adverbial will naturally be incompatible. Likewise there is a mismatch with an accomplishment and a durative adverbial. The durative adverbial picks out

---

[31]Originally coined in (Bennett & Partee, 1978).

an arbitrary endpoint of an activity, whereas an accomplishment has, built in, an actual endpoint that needs to be picked out in some other manner. More will be said of this in the following sections on event structure and coercion.

Again, the linguistic tests used do have some problems. First of all, there are the actual exceptions. For example, it is perfectly possible to use a *for x time* adverbial to indicate an accomplishment – *I ran for two hours – exactly as I intended.* This is a case where there is a natural, non-arbitrary endpoint associated with something that is normally considered an activity (i.e. *running*), but is treated as a situation with an inherent goal (i.e. an accomplishment). (Dahl, 1981) points out that English is not the best language for these types of tests:

> Some languages are more systematic than English in distinguishing indicators of actual and potential terminal points. Thus Swedish use different prepositions...

> (27) Jeg   reser till Frankrike *på* två månader.
>   I('m) going to France   for two months.

> (28) Jeg reste   i Frankrike *i*   två månader.
>   I   travelled in France   for two months.

> it is even possible to construct a sentence that contains both kinds of terminal point indicators...

> (29) Han   har suttit inne *på* två år   *i* sex veckor.
>   He('s) been sitting inside for two years for six weeks.
>   'He's been serving a two-year sentence for six weeks.'[32]

Another way to illustrate the telic/atelic distinction for dynamic situations is the entailment from the progressive. Commonly called the 'imperfective' paradox, it is a well known fact that the implications of a process-type aktionsart in the progressive differ profoundly depending on whether the aktionsart is an activity or an accomplishment. For example, compare the following two sentences:

(30)   a. John is pushing a cart.
        b. John is building a house.

If (30-a) is true, then the sentence *John pushed a cart* must also be true. This is not the case with (30-b), however. While it is certainly the case that (30-b) implies that *John has been building a house*, there is no guarantee that John

---

[32]The examples numbers have been modified from the originals. Also, the word-for-word gloss on (29) was added by this author for clarity.

actually finishes. Thus, another way to detect whether a process VP is either an activity or an accomplishment is to determine whether or not *X is ϕ-ing* entails that *X ϕ-ed*. If so, then the VP has, as default lexical entry, activity status.

Again, the notion of a goal reached, or object completed is crucial for understanding the difference between activities and accomplishments. As we shall see in more detail in the next section, an activity just consists of a 'uniform', durative, process. An accomplishment, on the other hand, consists of this as well, *plus* a 'goal' to be reached. The entailments from the progressive of the two aktionsarten seem less 'paradoxical' with this in mind. Roughly, if we take (Comrie, 1985b)'s view that the present tense merely says that a situation is true at the moment of speech (for non-narrative present), then the differing entailments fall out quite easily (informally at least).

If an accomplishment is viewed, broadly, as an activity plus a goal then it can readily be seen that, first, the present progressive picks out the 'activity' bit and says that it holds now.[33] Thus, the simple past rendition of the same activity would then hold as well. Observe that if (30-b) is true, then so is *John was building a house*. The difference, however, is that as of the time of speech of (30-b), the house (by definition, assuming that reaching the goal stops the activity) is not built. In fact a direct entailment of (30-b) (said more or less contemporaneously) is indeed

(31)   John has yet to build a house.

Recall that the present tense merely says that a situation holds at the moment of speech. Therefore, when it is applied to an accomplishment, we only are sure that the activity portion is true (and in process now). Certainly, there is a default inference to the effect that if John is building a house then eventually (barring catastrophe or ineptitude) a house will be built. But this is different from (classical) logical entailment.

The standard formulation of the progressive or imperfective 'paradox' is the different entailments that arise with past progressive instances of sentences with activities and accomplishments:

(32)   a. John was pushing a cart *entails* that John pushed a cart.
          b. John was building a house *does not entail* that John built a house.

Parallel to the present tense, the simple past tense merely requires that the situation hold at the event time (either overtly stated or contextual). Again, nothing in the tense coding says about what goes on before or after the event time, although certain aspectual markings such as the *habitual* would require this.

---

[33]The idea of the progressive placing the temporal perspective in the middle of an activity is more than a handy metaphor. (Bybee, Perkins, & Pagliuca, 1994, p. 127-130) has a survey of sources of the progressive, many of which are grammaticalised from words for *stay, sit, lie, be in a place*, etc.

Moreover, the progressive can only apply to an activity. Recall the infelicity of states with the progressive unless they are reinterpreted as activities. Likewise, the more punctual events discussed in the next section are also incompatible with the progressive unless suitably reinterpreted (compare #*The window is shattering* to *The windows are shattering*).

Comparing the two examples in (32), we know that the activity of pushing a cart only consists of one event-type – namely an activity. On the other hand, (32-b) consists of a building activity plus a goal of a finished house (which can be seen as a resulting state followed preceded by an instantaneous change – more on this later). But the progressive, then, must apply only to the building activity, which necessarily temporally precedes the goal being reached. So, the past progressive *was building a house* only asserts that the building activity occurred at a certain point in the past. Therefore the finished house is indeed not entailed. As with the present progressive formulation, this is not to say that there is no entailment whatsoever from the first part of (32-b).[34] Indeed we have the awkwardly formulated

(33)   John was building a house *entails* John was engaged in (house)-building activity.

that parallels the entailment of a past progressive activity. Moreover, we also know that John built at least part of a house. This can be seen by comparing the following:

(34)   a. John was building a house, but never finished it.
        b. ??John was building a house, but never even started on the foundation (i.e. didn't even build a bit of a house).

There is also, certainly, a default entailment that (unless told otherwise) the house is actually finished, but it should be noted that the other entailments (building activity engaged in, partial house built) are strict. It seems then that the 'paradox' really does not lie in the imperfective at all, and I would argue that the progressive treats activities and accomplishments uniformly. This is not to say that one can make the entire problem disappear, only that it lies elsewhere. I would say that this 'elsewhere' is in a proper treatment of telicity i.e. how to formally construct the telic goal so as to predict the correct strict and default entailments.

---

[34]A major difference between the present and past progressive forms is that the present form can be used where no *building* activity is taking place. With *John is building a fort, tomorrow*, his intention (or someone else's who can make him) is enough. This intention or plan leading up to an activity can in turn be seen as a preparation activity that is the case *now* thus justifying the use of the present progressive. From the past viewpoint, if no building activity ever occurred, one would much more likely say, *John was going to build a fort, but then thought better of it.*

Before moving on to the final major distinction, it is helpful to step back a bit and examine what these differences actually are. So far, we have seen the distinction between stative and dynamic situations as well as two distinct types of dynamic situations. At the end of the last sections we saw Comrie's view that the dividing line between what is a stative situation and what is a dynamic situations is to a certain extent rather slippery. While not totally arbitrary (e.g. one would not expect to see a language classifying *hit* as stative), there is a fair amount of 'wriggle room' in the way reality is partitioned. Again, the point needs to be made that the typology of event-types is *not* ontology. That is, it does not seek to find out the different kinds of events there are directly in the world. Rather, it is a typology of the different ways humans deal with things that happen in the world – the different way events are conceptualised.

Just as it may indeed be a matter of perspective whether or not a given situation is classified as stative or dynamic, it is also often the case whether or not something is classified as either an activity or an accomplishment. For example, the same bit of running can be seen merely as an activity or (if there is a specified distance or goal in mind) as an accomplishment. Thus both *I ran in the park today* or *I ran five kilometres in the park today* may describe exactly the same situation. A ship sailing on a certain trajectory may said either to be *sailing toward the North Pole* or *sailing to the North Pole*. It is easy to see from the progressive entailment test that the former is an activity while the latter is an accomplishment.[35]

As should be clear, it is often the case that way a particular situation is looked at is a matter of perspective and the speakers goals. States can be conceived as inherent or only temporary, in a more activity like way, or focus on the entry into the state rather than the state itself. Further examples of coercion and event reporting will be seen in this chapter and in a more psycholinguistic light in the next. In all cases, the deciding factor is how the speaker either perceives or wishes to present the event. This flexibility in conceiving and reporting events depending on what perspective the speaker is taking is only a symptom of the centrality of perspective taking in general. (Tomasello, 1999) amongst others stress the importance of perspective taking as part of a general cultural process that facilitates both the use and acquisition of language.

Instead it is the grammatical form that will often trigger the aktionsart into being either an activity or an accomplishment. This can be done with the count features on the object of the verb (*ate apples vs. ate two apples*), particles (*ate the apple* vs. *ate at the apple*), or even special types of quasi-auxiliaries called 'aspectualizers' – if John *finished* running rather than *stopped* running then he must have had some goal in mind (one hour, or running until the shoe leather

---

[35]It should be noted that in all the above cases, it is not the case that the two forms are semantically equivalent (in general the accomplishment implies the activity but not *vice versa*) only that these are two possible perspectives for describing the same bit of 'reality'.

falls off).

Sometimes, however, only context can determine the event-type. Notoriously, VP's such as *reading a book* or *writing an essay* can be used to describe either activities or accomplishments. Unfortunately, in this case, even the *stop* vs. *finish* test is unable to differentiate:

(35)   a. John stopped reading the book.
          b. John finished reading the book.
          c. John finished running a mile.
          d. ?John stopped running a mile.

Sentence (35-b) is an unequivocal accomplishment, where there is a clear goal that has been finished. (35-a) remains ambiguous – John could have given up the goal of reading the book (leaving the goal unrealised), or merely stopped reading the book for the evening, having no particular goal in mind with the reading.[36] With *running a mile*, it is much clearer – (35-c) entails that the goal was reached, while (35-d) is rather strange. The aspectualizer *stop* most usually applies to activities – hence the infelicity with *run a mile*. The best that can be done to get at the intended meaning of (35-d) is to say something like *John stopped running before he even ran the required mile.*

Achievements and Points

The two dynamic eventualities discussed thus far have been activities and accomplishments. Activities, which are extended in time, dynamic, but have no inherent endpoint. Accomplishments, which share both temporal extension and dynamism with their fellow processes have an additional feature of telicity – that is, a natural endpoint (i.e. a goal reached, something created, something destroyed, etc.) At a certain point in the unfolding of the accomplishment there is a sort of 'abrupt' change that contrasts with the 'gradual' change of activities. One moment (conceptually, at least), a builder is busily constructing a house and the next moment there is a house completed. It seems that (in ordinary English, at least), this sort of scenario is what people commonly mean when they use the term event. Something *happens* (unlike with states) rather quickly, leading to some sort of change.

For more concrete examples, observe the following data:[37]

---

[36]It is probably the case that Dahl has a valid point in noting that English is not the best language to test terminal points. As already mentioned, the various tests are not infallible. Indeed, when examining various experiments on children regarding event completion and telicity in the following chapter, it turns out that intuitions on the telicity of eating and drinking predicates are the among the least consistent.

[37]Examples adapted from (Croft, 1998).

(36)   a. John shattered the window.
         b. #John shatters the window.
         c. #John is shattering the window.[38]
         d. John is shattering windows.

(36-a) can be considered an archetypal achievement, in that a change occurs instantaneously. For this reason it is then straightforward that (36-b) and (c) would be rather odd. Assuming (36-b) is used in the standard (not narrative) present tense, there just isn't enough time to report the act as it is happening. Rather, one must say that John just shattered the window, or that he is about to do so. Similarly, the progressive use in (36-c) is anomalous due to the 'durative' (in progress) nature of the progressive. But then, why is there no # by (36-d)? As seen before, the object of the VP makes all the difference. Here, John is shattering a number of windows, and what is being 'progressivized' is precisely this ongoing activity.

As noted by Comrie, another punctual verb, *cough* has but one interpretation in the progressive (*He was coughing*) – a number of coughs in a row. In fact, when a verb such as *cough* is used in a language with morphological imperfective marking (e.g. French) the interpretation of a sentence like *il toussait* cannot mean that only one cough happened.

Of course, it can be argued that virtually nothing happens strictly instantaneously – a cough takes time and can probably be broken down into various phases by a respiratory specialist. Likewise, the shattering of a window is probably quite a complex event when viewed at 1/1000 time. And again, the answer is that linguistic distinctions and conceptions need not strictly correspond to 'objective' reality. From (Comrie, 1976a) (p.43):

> For as long as we are in real time, it is unlikely that anyone would want to refer to the duration or successive phases that make up this situation

Comrie then goes on to develop a scenario where, with slow motion film, a lecturer in anatomy can felicitously utter *and now the subject is coughing*, where only one cough is ever coughed. Despite this contrived example, there is good reason to think that this sort of 'punctuality' is a good candidate for an linguistic/aspectual category. Both Russian and Hungarian have classes of verbs with particular suffixes that can only be viewed as either punctual or iterative. Indeed the Russian verbs for 'cough' *kašljanut* and 'flash' *blesnut* only appear in the perfective aspect.

So far, all of the examples we have seen of achievements are those such as *cough, flash, shatter, etc.* that while indeed perceived as punctual can be, with

---

[38]There is a futurate progressive reading that is perfectly felicitous, i.e. *John ripped up the floor today and he is shattering the window tomorrow*, where what is in progress is the plan to shatter the window.

some context-stretching, made to seem more 'durative'. The question then raises itself as to whether there are actually any situations that might be strictly punctual. And in fact there are. Consider *win a race* or *reach the summit*. These can, in a sense, be considered more as conceptual situations than the quite physical (but very short in duration) ones listed above. Conceptually at least, it is easy to imagine that there is a moment before the subject wins the race or reaches the top and another moment when they have with no intervening time. That is, unlike a *cough*, or a *flash*, there is no way to further break the situation down into phases.

But with these truly punctual types of situations, an interesting thing happens when they are put into the progressive.[39]

(37)   a. Shari is winning the race.
         b. Bill is reaching the top.

In both cases, there is no iterative interpretation as with *coughing*, but rather a reference to the time just before the achievement is reached.[40] That is, it becomes something akin to an accomplishment, where the progressive picks out the run-up activity that will (if everything goes well) lead to the goal being reached. For this reason, VP's such as *win the race* and *reach the top* are often considered to be 'gradual' achievements, while VP's such as *flash* and *cough* are 'points'.

Another way to see the difference is to go back to the container/durative adverbial tests:

(38)   a. John reached the summit in just two days.
         b. It took Marcia ten minutes to find her keys.
         c. #John reached the summit for two days.
         d. #Marcia found her keys for ten minutes.

(39)   a. The light flashed for five minutes.
         b. Billy coughed for three minutes straight.
         c. ?The light flashed in five minutes.
         d. ?Billy coughed in three minutes.

With the gradual achievements, the container adverbial serves to profile the time leading up to the achievement. And while it was noted earlier that there is a similarity to accomplishments, there is, however, one major difference. With a prototypical accomplishment such as 'build a house', the associated activity leading to a built-house goal is, indeed, building. This is different for the gradual achievements. Even though it is fine for many English speakers to say that some is

---

[39]This might be seen as a contentious bit of English data, but it is fine for myself and many other speakers. Comrie begs to differ.

[40]Ignoring the other use of *is winning* that could be used for a runner who is leading even if the race has just started.

reaching the summit, the associated activity is not *reaching*, but rather something like *climbing* or *hiking*. And in fact, it is absurd to say that *Marcia is finding her keys*. Rather, she is *looking* for them.

Combining durative adverbials with 'points', what is profiled is the time that the iteration punctual events lasts. Container adverbials are not quite as felicitous. I have put '?'s next to (39-c) and (d), as enough contextual flexibility does make them interpretable. Imagine Billy being examined by a doctor and told to cough, but Billy is rather shy. In this case one could imagine (39-d) or at least the sentence *It took Billy three minutes to cough.*

Binnick uses a few tests to distinguish between accomplishments and achievements. (Binnick, 1991, p.177) gives both adverbial and entailment tests, some of which are unfortunately not infallible. The first claim is that achievements are felicitous with container adverbials (*in x time*), but not compatible with durative adverbials (*for x time*), which is shown in (38). He gives the infelicitous example

(40)   #John noticed the painting for a few minutes.

But with the achievement *win the race* this is perfectly acceptable.

(41)   John was winning the race for a few minutes, but then fell behind.[41]

His second distinction is that achievements are generally not compatible as complements of *finish*. His examples are

(42)   a. *John finished noticing the picture.
         b. John finished painting the picture.

As noticing is generally conceived as punctual, it is rather hard to conceive of it being begun or finished.[42] Moreover, even with an achievement that can be conceived as being more drawn out like *win a race, John finished winning the race* is certainly anomalous.

Finally, there is an ambiguity with *almost* used with accomplishments that doesn't occur with achievements. Take the following two sentences.

(43)   a. John almost noticed the painting.
         b. John almost painted the picture.

Sentence (43-a) entails that John did not notice the painting, while (43-b) can mean that John never even engaged in picture-painting activity or just that he didn't finish. However, this seems to elicit more the difficulty in assigning a

---

[41]A partial reason for this discrepancy is given in the section on Event Structure. For now, it is enough to say that *win the race* in the last example is used as an accomplishment, not an achievement. The ostensibly similar *reach the top* patterns with notice in being normally unacceptable with durative adverbials.

[42]He notes that *John stopped noticing the picture* has a habitual reading. For example, once one has a picture in the house for a while, one may stop noticing it.

strict telicity to what are normally considered accomplishment VP's. The first interpretation of (43-b) construes the VP only in its activity phase, while the latter interpretation takes the goal into account as well.

At least, even the less well behaved achievements such as *John almost won the race* needn't even entail that at any point *John was winning the race*. It might have been the case that John ran the entire race as a very close second, but could never pull ahead.

The major point to keep in mind regarding the fallibility of the various tests and the slight variance between difference concrete examples of a certain Aktionsart (e.g *notice a painting* vs. *win a race*) is that the event typology attempts to hang a variety of situations on a simple four or five part distinction. But all is not lost. In Section 2 we will see a general notion of event structure that the aktionsarten can be seen as participating in. Secondly, this notion of a general event structure allows for the possibility of various construals that defy the default mapping, but provide a principled reason to account for the what have appeared to be failures of the various tests. Before this, we shall examine a final refinement of aktionsarten.

Series

So far we have seen what may be called the standard 'Vendlerian' event types (States, Activities, Accomplishments and Achievements) plus an extra refinement, the splitting of Achievements into gradual achievements and points. Before moving on to the section on event structure and coercion, it is useful to look at one more proposed event type – Series, first discussed by (Freed, 1979).

(44)   a. John kept being ill. (series)
        b. John was ill. (state)
        c. John understood French (state)
        d. ?John kept understanding French. (attempted series)
        e. John loved Mary. (state)
        f. ?John kept loving Mary. (attempted series)

Ostensibly these seem to be akin to iteration, but there are some significant differences. Moreover, it also may be seen as a sort of 'meta' event-type, with a rather close relation to genericity. I find this a rather difficult category to get a firm grasp upon as, unlike the Vendlerian types, there seems to be no default relation between a VP and a series. Rather the tests for a series rely on both compatibility with certain aspectualizing verbs as well as the presupposition and entailment properties of individual sentences. It might help to see some of the data that Freed uses. For example, series, in contrast to states, allow the aspectualizer *keep* with the gerund (i.e. *keep* VERB-*ing*).

To me, (44-a)'s most natural interpretation is not that John was ill for five days straight and just couldn't get better (this seems better expressed by the stative *John continued to be ill*) but rather that John was ill on a number of successive occasions over an extended period of time that also consists of many intervals where John was not ill at all. This differs from the iteration seen in the last section (with *points*) in that an iteration of coughs, while consisting of a number of coughs is intuitively seen as a single event.

In contrast, a series of illness states seems less of a unitary event, but rather a number of 'separate' events bundled up together into a new one. I have labelled the examples with '?'s as *attempted* series in that with a little context they are interpretable – but then, as a series of understanding or loving events. For example, imagine that John is unconfident about his French language skills and convinces himself that he does not understand French. Nevertheless, every time a Parisian speaks to him, he understands.

With the concatenation *stopped* VERB-*ing*, we can see a contrast between series built upon activities or achievements and activities and achievements proper based on presupposition and entailment relations. (Freed, 1979)(p.59):

> Series: *John stopped driving to work* presupposes *John used to drive to work* and entails that *John does not (is not) drive (driving) to work now*.

> Activity: *John stopped running* (on one occasion) does not presuppose that *John used to run*, nor does it entail that *John does not run (is not running) now*.

> Achievement: *John stopped noticing the lint on his suit* (on one occasion) does not presuppose that *John used to notice the lint on his suit*, nor does it entail that *John does not notice (is not noticing) the lint on his suit now*.

Note that both sentences in the activity and achievement parts of the above quote indeed do have interpretations as series (in this case taking on a rather habitual reading), and it is generally the case that only context (or something explicit like her use of 'on one occasion') makes the interpretation clear. For example, John's noticing lint on his suit may have been a debilitating compulsive behaviour that is finally stopped after years of therapy. This can be viewed as a sort of generic series, as there is no reference to any individual noticing event *per se*. The example where John keeps being ill, can, as a series, have either a generic or serial (these are the two types of Freed's series) interpretation. The serial one can best be summed up as *John keeps being ill. It's getting rather annoying – at least once every two weeks there's something amiss...* Another example from Freed (of a non-generic series) is the contrast between the state *she slept from*

noon to 3:00 p.m. and *she slept in hotel rooms from May to August*, which is a series of individual events gathered together.

Based on some of the tests (especially the presupposition and entailment), there is certainly an argument to allow for yet another event-type (though derivative on the basic ones). What makes the series is a different sort of operation than, say, the iterativity on points or achievements. It may even may it easier to incorporate the notion of genericity or habituality into the framework in a nice way. Nevertheless, I think it safe to rule Series out of a notion of 'primitive' event-type, which I would restrict to the classic Vendlerian four and the points.

With these types in mind, the next section will look at a notion of event structure (taken mainly from (Moens & Steedman, 1988) and (Freed, 1979)), which will attempt to put together the various types into a unified whole. Intimately related to this is the notion of aspectual *coercion* which has been mentioned a few times earlier. And with a notion of event structure and coercion, we will wrap the section up with a look at how combined they give a window into the relation between linguistic meaning, context, and event interpretation.

2.2 Event Structure and Coercion

2.2.1 Introduction

Recall an example from Kalam in (9), repeated here for convenience:

(45) kotp gy,   np ñnp   gspyn
      house having-built-SS you intending-to-give-SS I-am-doing
      I am building a house for you

Here is an example of an accomplishment event type, put in the progressive. For our purposes, the interesting bit here is that the accomplishment is spread over separate clauses. One clause is the indication (or speculation at this point) of a house having been built. This is the clause with *gy* ('having built'), which is the generic activity/doing verb *g-* concatenated with a perfect-marker (*y*).That is, the building activity (which is linguistically represented by *g-* and the house (*kotp*) as object) has led to a state – that of a house having been built. Furthermore, a separate clause is used to indicate that the activity associated with the building of a house (i.e *building*) is going on now – seen in the sentence as *gspyn* ('I am doing').

That this is so explicitly spelled out in the language is rather uncanny as it strongly echoes the received wisdom on the temporal contour of events. In comparing the different event types, the connection between accomplishments with activities and achievements has been hinted at. They seem to share something in common with both. As it has already been demonstrated that often the difference

between activities and accomplishments could be merely that of perspective (e.g. the same situation could be described as either *running* or *running a mile*). An accomplishment, in a sense, could be seen as an activity leading to an inherent end-point, or goal. However one wants to characterise it, there is an activity at the heart of an accomplishment.

But then what of the goal aspect? Again, the similarity between accomplishments and gradual achievements (such as *reach the summit*) was noted. With both, a goal is reached (which we may think of as the 'event') leading to a consequent state (having built a house or having reached the summit). So, we can think of accomplishments as, in fact, something akin to an activity with a punctual achievement added in for good measure. And while Kalam speakers seem to have been rather advanced event semanticists, it took the primitive western semanticists until the 1970's and 80's to figure this out. Nevertheless, there are a number of similarly oriented notion of a general event structure that can be used both to unify the basic event typology, but also shed light into the various coercion phenomena.

2.2.2   Event Structure

In this chapter we have seen a more or less standard typology for different types of events. This phenomena is traditionally referred to as *lexical aspect*, as lexical VP's are commonly seen to denote one of the event types. However, one could take a more holistic view of event representation and meaning in general. On this view, there is no strict separation of the different event types. Rather, given the proper perspective one could view something conventionally thought of as an achievement as, say, an activity. Moreover, the lexical VP's that are associated with a default event type are allowed to shift to another given the right context or grammatical construction. But to do this, a notion of event structure that can subsume all the 'primitive' event types to allow such shift is needed.

Luckily we do have all the necessary ingredients.[43] At the end of the preceding section we saw how one could look at an accomplishment as a telic activity along with an achievement-like event that ends the activity and a consequent state. That is, there does seem to be a way to relate the various event types with one another. Though deceptively simple, the 'model' proposed by (Moens & Steedman, 1988) is quite a good start at this sort of unification. Their structure is called an event *nucleus*, and is rather closely related to an accomplishment. It can be visualised in the following:

---

[43]As with the rest of this chapter, I do not mean any specific formal machinery, but rather am sticking to the somewhat hazy realm of intuitive notions and pictures. Formalities will come later.

=============|/////////////////////////////

preparation | consequent
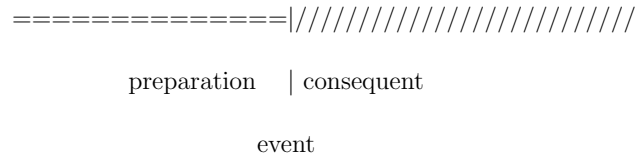
event

Keeping the stock 'building a house' example in mind, the preparation part of the event structure corresponds to the activity part of the accomplishment (or in (45) 'I am doing'). The *event* (proper) can be seen as the achievement part where the house goes from unfinished to finished. The part labelled 'consequent' can be seen as the other clause of (45), namely expressed in the English perfect as 'having built a house'. More generally, we can view the accomplishment as the most complex event type, with the others partaking of one or more parts of it. Thus, in general, the preparation of the nucleus can be seen to correspond to an activity. Recall that while atelic by default, any activity can in principle be made telic (even if the goal is merely doing the activity for a given amount of time). Similarly, we can take an achievement as the event part plus the consequent (for example, if I have reached the top, then I am in a state of having reached the top). Recall that the preparatory part of an achievement can also be focussed upon, if desired. Points can be seen as the punctual middle without the surrounding preparation or consequent.

One may want to make slight modifications, however. First, it is convenient also to take the telic goal directly into account. Note on this structure that there is only a preparation leading toward a goal, the *event* of the goal being reached and the consequent state. The addition of the reified object (not merely the intended end-state) is useful for dealing with the problems of telicity such as the imperfective paradox discussed earlier.[44] Also, in order to handle *ingressive* aspect for activities and accomplishments, it is important to realise that we can also look to what happens before the preparatory activity starts. Imagine, for example

(46)   John started building the house.

The complement is an ordinary accomplishment that has the structure sketched above. But the aspectual verb *start* is an event that triggers the preparation. However, this deceptively simple concept does account for many complex, hierarchical event structures.

This is because the above picture is meant to be skeletal. In fact, each of the parts can be iterated, depending on what level of granularity a particular event is viewed at. For example, this simple structure is perfectly fine for capturing the rather high-level notion of 'building a house'. But, of course, the preparatory building activity is not one activity, but rather many activities, accomplishments,

---

[44]The interlude sketches how to do this.

etc. that are conceptually thought of as activity leading to a house being created. Thus, the preparation bit can be seen as (amongst others) laying a foundation, putting up walls, laying plumbing, electrical wiring, etc. And each of these can be further divided hierarchically into more fine-grained descriptions. To quote Moens (2000) about the components:

> Each of them may itself be compound. For example, the preparation may be an iteration of some kind, the consequent state may identify a chain of consequences, and the core event may itself be a complex event, such as an accomplishment.

For example, the futurate progressive, *John is flying to London tomorrow* demonstrates how the core event may be complex. The preparation is not flying activity as it would be in a present progressive, but rather John's plan to go. The event that is then triggered by the preparation is not a punctual event of arriving in London, but the temporally more complex (and extended) accomplishment. But couched in the event structure as perfectively viewed, with the imperfective indicated by the progressive marking denoting the planning preparation.

In fact, but there is both linguistic and psycho-linguistic evidence for this sort of hierarchical event segmentation. In the next chapter we will discuss experiments by Zacks and Tversky[45] that indicate that certain event-types such as 'building a house' or 'going to a restaurant' are stored in the mind in such a segmented and hierarchical fashion. Going back once again to Kalam, this segmented, hierarchical structure is often encoded directly into the grammar. We have already seen in (45) that (at least with a beneficiary) the accomplishment of 'building a house' must be spread across multiple clauses that roughly correspond to the above event structure. But the connections go deeper than this.

(Pawley, 1987) shows in great detail how Kalam speakers rely on a schematic event sequence to describe complex events. For example, in reporting intentional actions, the schema is as follows:

1: Movement to Scene of First Action
2: Action (iter.)
3: Movement from Scene of 2 to Present or Final Scene
4: Action(s) at Present or Final Scene (iter.)

Each part of the sequence may be internally complex, and 2 and 4 may even contain an event-sequence themselves. But the typical use is a serial verb sequence with no nesting. This can be seen with a Kalam report of a typical and simple 'complex' event such as gathering firewood:[46]

---

[45]See (Zacks & Tversky, 2001) for a summary of a variety of related experiments.

[46]The use of the '/' in the Kalam text is only to delineate the separate scenes and is not part of the language.

(47)   am/ mon pk/   d   ap/   ay-
        go   wood strike get come put

So, what would normally be expressed in English as a simple verb phrase (*gather firewood*) is instead expressed as a sequence of simple events – the first scene has the actor moving to the wood location, in the second scene the cutting is done, the third scene has the cutter bringing the cut wood home, with the final scene being the action of putting the wood in its proper place. In terms of event structure, we can envisage the 'go' to lead to the inceptive event, the 'wood, strike, get' as the preparatory activity (iterated, of course) and the final action of 'put' to be the terminating event.

This hierarchical goal/plan/event structure will come back in more detail in the chapter about psycho-linguistics, but it is useful to make a few, general comments here. To quote Zacks and Tversky's plausible speculation:

> The most important legacy of the evolutionary forces shaping humans' cognitive architecture is probably a general mechanism or set of mechanisms for inferring and reasoning about causes and goals.

The upshot of this is that production and comprehension of language about events draws on these very mechanisms. More importantly for the purposes of the interaction of aspect and tense is the claim ((Steedman, n.d.), (Lambalgen & Hamm, 2005)) that the cognition of time itself may be parasitic upon goal/plan mechanisms. More concretely, for Moens and Steedman, in a tensed accomplishment like 'Harry climbed to the top' the temporal reference can actually be used to refer to a default relation between an antecedent (climbing) and goal (reaching the top). Also, the temporal profile comes along with the notion of a plan/goal relationship (as executing a plan toward a goal takes time). Similarly, many other aspectual and temporal phenomena can be see as the cognitive representation of goals and sub-goals.

Moving back to the matters at hand in this chapter, an integrated part-whole event structure can also have something rather important to say about coercion phenomena, and more importantly, meaning in general. We shall see this in the next section.

2.2.3   Coercion

Until now (though with the proper caveat), the talk has been of, e.g., stative or accomplishment VP's. The standard diagnostic tests have been given to illustrate, for example, why a particular VP should be considered stative rather than as an achievement or an accomplishment. Some authors (e.g. (Olsen & Weinberg, 1999)) even put the lexical aspectual type of the V or VP as lexical features

of said V or VP in a generative framework.[47] However, matters are slightly more complex. It is important to different aspectual classes of predicates in a particular language from the aspectual types which can be seen as representations of semantic structures. One of the reasons is that there is not a simple (1-1) mapping from predicates to event-types. Instead, one should think of 'stative VP's' or 'achievement VP's' rather as a default as to how a situation is generally construed. But, being a default, context, or a particular linguistic construction can change the default event-type associated with a given VP to a different event-type. This is what is called coercion.

Achievements and Points

We have already seen an example of coercion before, when comparing the differences between Achievements and Points. When put in the progressive, they may behave differently.[48]

(48)   a. John is reaching the top.
         b. John is coughing.

(48-a) exhibits the change from an Achievement to an Accomplishment, or what have been termed as 'run-up'-Achievements. As the name alludes to, there is a shift in focus from event and post-state part of the event nucleus (typical of an achievement) to the activity that is leading up to the achievement. Essentially, this is the addition of the preparatory activity to the given event structure of the achievement.

   (48-b) is slightly more complicated. Recall that a point is merely the instantaneous event part of the event nucleus. The application of the present iterates the coughing into an activity. Just as the preparation of the 'house-building' could be seen to be a hierarchical structure of various other events, the activity here will consist of an iterated series of coughs, taken to be a uniform activity.

   The classical entailment test also show that the (48-a) is a progressivized accomplishment, and the latter a progressivized activity. Notice that if a climber is reaching the top, there is no strict entailment to the fact that the climber will actually reach the top.[49] In contrast, if John was coughing then there is certainly the entailment that *John coughed*.

   Both examples can be viewed as the progressive construction systematically changing the default aspectual type of the VP it applies to in order to remain interpretable. As the progressive needs to apply to an activity, an activity is

---

[47]Verbs are not marked *state* or *accomplishment* as such but reduced to features such as **+dynamic, +telic**, which are somehow innate.

[48]That is, if they are grammatically acceptable. This will in some more marginal cases vary among speakers.

[49]Or to put in terms of the progressive paradox. *John was reaching the top* does not imply that John did reach the top.

added in the first case to yield an accomplishment, and the progressive proceeds as normal. In the second, the simple addition of an activity would yield an event structure with an activity and culminating event, but no end-state. In English at least, this type of event does not seem to exist.[50] So instead, there is move to iterate the instantaneous event and construe it as an activity to then apply to. Note that a point with progressive marking does seem to entail such an iteration. If John was coughing, or a light was flashing, then both of these events must have occurred more than once.

Matters can be complicated a bit more by the addition of other constructions to the given lexical material. For example, there is the (hopefully metaphorical) VP *cough up a lung*, the coercion of a point to an achievement with the addition of a rather painful and disgusting end-state. The addition of the progressive (*He's coughing up a lung*), yields an unpleasant run-up achievement.

Durative (as opposed to punctual) temporal adverbials force an iterative reading with achievements. For example, observe the well-known following pair (from (Moens & Steedman, 1988)):

(49)   a. *John arrived all night.
        b. Visitors arrived all night.

The adverbial *all night* forces a reading of multiple achievements over the specified time. Sentence (49-a) is bad precisely because of the incongruity of the situation. In contrast, visitors arriving separately over the period has no such incongruity. Similarly, changing the object of the achievement to a plural, allows for the iterated, activity reading in the progressive. For example, *John is now reaching the summit of every peak he attempts* describes an ongoing activity (or progressivised habitual).[51]

The above gives us a few ways that in the context of certain grammatical constructions (progressive, adverbials, even change in the semantic status of the argument) that VP's that are associated with either achievements or points can, in these particular contexts, undergo a shift to another. However, as can be already seen by (49-a), not everything goes here, but perhaps more than is often recognised in the literature. (Comrie, 1976a) rejects the use of 'John is reaching the top' for a run-up achievement, although many speakers find this form perfectly acceptable and it is well-used. Aside from speaker variation, there are certain achievements that do seem rather difficult to turn into run-up achievements.

(50)   a. *Mary is finding her keys.
        b. *Julie is noticing the lint on her blouse.
        c. Mary is finding a flaw in everyone at the party.

---

[50]Similarly, the perfect (*has coughed*) yields just this end state.

[51]Notice that here we can replace the progressive with *reaches* felicitously. *John reaches the top*, however is no longer a run-up achievement, but rather a completive, narrative present.

    d. My mom was noticing that my zits were freshly squeezed all the time. She begged me to stop.[52]

    e. Slean is finding her place among other great Canadian singer-songwriters. There is no place, yet, where she ought to be – she's on her way there, wherever "there" is.[53]

Concrete uses of *find* seem impossible as a run-up achievement (*looking for* would be used), but are quite common in more metaphorical uses such as (50-e). The *way* construction (see (Goldberg, 1995)) is also quite commonly used with *find*. The interesting bit about the last example is that it shows that even in this more metaphorical use, it definitely has the entailment properties associated with run-up achievements and accomplishments. Of course, (50-c) is the expected iterative reading with a plural object.[54]

    Sentence (50-d) is a typical example of *notice* in the progressive. For the most part, these types do not seem to be run-up achievements or accomplishments but an extended use of the progressive (perhaps for politeness), such as an applicant for a job saying "I was just noticing your advert and..." But, the above do have a gradual, process flavour. However, they too, have a strict implication to the goal.[55]

    Finally, it should be pointed out that even points, given correct context, can have a run-up achievement reading, rather than the standard cyclical/iterative one. (Croft, 1998) gives an example of *the light is flashing* in the situation of an observer viewing a lighthouse from a far distance and the light turning around at a slow speed. Interestingly, the light is not really flashing at all, but would give the appearance of a process building up to a flash when the light comes round again. From the speaker/observer's perspective, this is enough to justify the reading.

States

In the earlier sections it was observed that often the boundary between a state and an activity is a matter of perspective. Applying the progressive to states[56] shows this in two slightly different ways:

(51)   a. Sylvia resembles her mother.
        b. Sylvia is resembling her mother more and more everyday.
        c. Jeff is a jerk.
        d. Jeff is being a jerk again.

---

[52]This example is from facepick.tripod.com/confess.htm.

[53]www.rambles.net/slean universe.html.

[54]There may also be a run-up achievement reading for this as well.

[55]This is not unique – *is dying* and *persuade*, to cite just two, have similar properties. See (Comrie, 1976a).

[56]Examples from (Croft, 1998).

(51-a) is a prototypical stative relation, but in a progressive construction as in (51-b), this relation is looked at with more granularity. Instead it is an activity that is seen to be increasing (in the strength of relation) over time. Nevertheless, it is still an activity, as it certainly is the case that (b) does indeed imply (a). The relation (of resemblance) is presupposed, and what is asserted is that the relation is changing in a certain direction. Of course, this goes against the (prototypical) notion that states are unchanging; however, as the we must coerce the state to a gradable activity via the progressive construction, the initial intuition can be maintained.

The contrast in (51-c) and (d) is a bit different. Unlike (51-b), (d) is an undirected, unmetered activity. It is not asserted that Jeff is being more of a jerk by the minute (metered by the consumption of Tequila), but rather his current behaviour (actions) are typical of someone who is a jerk. This also contrasts with (c), which asserts that 'being a jerk' is something inherent in John's nature (a prototypical stative reading).

Importantly, there is no implication from the disposition[57](51-d) to (51-c). While with the adverb, *again*, this may seem a reasonable conclusion, it would be strange to conclude from the sarcastic, *John is being smart, for once* that John is intelligent in a more general way. These sorts of dispositional constructions are an example of the blurry boundaries between states and certain types of activities.

Ingressives and Inchoatives

As a final set of coercion examples,[58] ingressives and inchoatives profile the part of the event structure that leads into event-type of the VP in question.[59] Often, a container adverbial construction is enough to trigger this coercion:

(52)   a. In two years, she was president of the company.
       b. The horse was galloping in two minutes.
       c. The light flashed in a few seconds.

Recall that in the earlier sections on states, activities and points it was noted that container adverbials are incompatible with such event types (being more associated with either achievements or accomplishments). This is still the case and precisely the point. While the VP's above are states, activities, and points respectively, the sentences they occur in are not, but rather resemble something akin to an achievement (with the preparatory part also profiled).[60]

---

[57]This term for a coerced state to activity is from (Croft, 2000b).

[58]I mean final for the purposes here. Various coercion phenomena both on the event-type and temporal level are ubiquitous and too numerous to list.

[59]Inchoatives describe the entry point into a state, ingressives the beginning of an activity.

[60]These examples, along with earlier ones, support (Freed, 1979)'s claim that Aktionsart must ultimately be determined not from the verb or even VP, but the entire sentence (in context).

Keeping the event nucleus picture in mind, we can see the container adverbial profiling the preparatory activity (whatever that may be) leading to a certain event that causes a change of state (i.e. not president to president, etc.). Again, having the entire event structure at our disposal (rather than only the bits standardly denoted by the VP's), makes the interpretation quite easy to compute.

Before moving on to ingressives, it is helpful to see that this type of inchoative coercion is not merely a quirk of English or other western languages (such as the well-known French examples), but is rather a common phenomenon. Take the Micronesian language, Palau.[61] The verb for English *go*, *mo*, has a typical use as a futurate marker,[62] but in a certain inflectional construction, can also coerce certain states into inchoative readings (semantically this is really not so surprising):

(53)   a. Ak mo me saab er a klukuk
         I'm going to study tomorrow.
         b. A Toki a smecher.
         Toki is sick.
         c. A Toki a mlo smecher.
         Toki got sick.

Again, the entry into the state rather than the state is profiled in (53-c).

Ingressive aspect also profiles the beginning of or entry into a situation. In English, this is done with the 'aspectualizers', such as *begin, start, commence*, all with subtly different properties (and dependant on whether an infinitival or gerund is used).

(54)   a. Lillian started studying for exams last week.
         b. Lillian started to study, but decided to go to a movie instead.
         c. Lillian began to study, but decided to go to a movie instead.

(54-a) and (b) have subtly different properties. Sentence (a) profiles the beginning of the activity, but nevertheless implies that the activity has indeed occurred. With *start* and the infinitival, there is no such implication. Instead, we have a more complicated coercion. What is profiled, instead, is the preparation that, if successful, leads to the activity being initiated. (54-b) could be a situation where Lillian didn't get one bit of studying done, but merely cleaned the table, sharpened her pencils and took her books out of the book bag. If we imagine 'events' as chains of initiating and terminating events, surrounded by states and activities, this is easy to understand. The 'started' in (b) refers to the initiating event that drives the preparation to the initiating of studying event, not the studying event. Note that the periphrastic 'about to' accomplishes the same

---

[61](Josephs, 1975).

[62]See (Bybee et al., 1994) and (Hopper & Traugott, 1993) for a comprehensive survey of this sort of thing in numerous non-genetically related languages.

effect. Palau has a similar phenomenon, but is again done by inflection on the verb:[63]

(55)   a. A Droteo a *melekingu* er a demal a toki el kirel a chebechiielir.
       Droteo is about to talk to Toki's father about their marriage.
       b. A Droteo a *melekinga* er a demal a toki el kirel a chebechiielir.
       Droteo has started talking to Toki's father about their marriage.

(a) (at least the gloss) seems compatible with a situation where Droteo actually is in front of the potential bride's father, and open's his mouth, but no words manage to come out, which could easily be described by *started to talk.*

Conclusion

At the beginning of this chapter, the standard event typology was reviewed, with different VP's corresponding to types of events with very different properties (e.g. durativity vs. pointlike, telic vs. atelic, etc.). Standard contextual tests were used to determine different classes of VP's on the basis of their grammaticality/acceptability in these tests. But by the last section, we have seen that much more is acceptable than was initially thought. Indeed, there is no simple, inflexible mapping from a given class of VP's to a semantic event-type. Instead an underlying event structure (such as the Moens and Steedman event nucleus) was shown to serve as the foundation allowing for systematic changes in the default meanings of a given aspectual class. This has been said in better form in (Croft, 1998)):

> The observations...lead us to two conclusions. The first is that the semantic representation of an event denoted by a verb must be complex, involving not only what we normally think of as "what the verb means" but also the event leading up to it (the run-up process or preceding state) and the event leading out of it (resulting state or, in the case of cyclic achievements, the return to the original state). Hence we must be able to specify what part of the event is actually described in the sentence and what is in the background.

> The second conclusion is that some adverbs and aspects (*suddenly, for a week, in two hours*) like the simple present/progressive distinction involve a construal of the event, specifically some particular time interval; but which time interval is profiled depends on the temporal structure of the event. The container adverbial profiles the process leading up to the last natural transition point in the event frame – the endpoint for telic events and the inception for atelic events (including states, inceptive states and achievements). the durative adverbial

---

[63]From (Josephs, 1975).

profiles an event continuing beyond the first transition point in the
event frame–that is, the inception for all event types, including telic
events.

Indeed, it is this background that allows different lexical or grammatical constructions to shift the meaning in a way that is quite easy to interpret for the hearer. This gives a picture of meaning that is not fixed by features on a given lexical entry and projected into the sentence upwards, but rather the combination and unification of (ostensibly) conflicting constructions to give rise to shifted meaning.[64] Ultimately, nearly anything is possible as long as there is enough context to allow for semantic agreement.[65] Of course, speaker acceptance will vary, but is certainly much wider than the standard literature would have us believe.[66] The reader may have found some of the coercion examples to be rather ordinary, and some rather strange. The more ordinary examples have acquired a conventional status and may hardly 'feel' like coercion anymore, but most likely began their existence as a novel example, that for one reason or another, became a standard facet of the language.

In the next chapter, we shall look at a good deal of psycho-linguistic literature and experiments that seem to support this more broad approach to events as evidenced by coercion phenomena and its interaction with event structure.

---

[64]The spirit of this approach to meaning (and also syntax) has been taking up recently in various flavours of construction grammars (Fillmore and Kay, Croft, Goldberg).

[65]Recall that it is rather hard to give a decent context to *John arrived all night*, but perhaps it is possible.

[66]On a related note, the reader is referred to (Bresnan & Nikitina, 2003) for new data on the dative alternation in English that confounds earlier approaches that relied on lexical class membership. By examining a large number of corpora including ordinary peoples 'fan fiction' on Internet web-sites, they discovered numerous counter-examples to what was an accepted data set.

Chapter 3

The Psycholinguistic Turn

In the preceding chapter the notion of event was examined from a semantic point of view. More specifically, we saw the standard event-typology, a structure on which the different event-types can be situated, as well as some ways language can encode these various event types. We also saw the way different grammatical constructions or lexical items can change the semantics of a particular VP to a different event-type in regular ways.

In this chapter we shall look at events from a psychological perspective – that is, how we cognitively represent events both in memory and as experienced, as well as how this manifests itself in language (i.e. event reporting). It turns out that many factors noted by psychologists as central to the conception of events were already seen in the semantic descriptions of events in the last chapter. These include notions such as intention, goal, and hierarchical representation.

3.1 Introduction

Take the following two sentence discourse

(1)   Max fell. John pushed him.

The most prominent reading of this is that the pushing event is the reason that Max fell.[1] The reader of this example will undoubtedly come up with this reading automatically. Indeed, as discourses go, this does seem to be a fairly uninteresting one.

Yet, there is a good deal that must be done to reach the causal interpretation. First, the reader/listener is presented with two events, both in the past (given the verbal morphology). Moreover, there must be a way of relating them to each

---

[1]There is the other, linear order of events reading, but this would normally have some added linguistic information, such as *Then John pushed him.*

other temporally (and to possibly a more general context). The question then arises as to how this is done.

Examining example (1), we can see that in the interpretation we are looking at, the temporal order of the events is opposite to the way they are presented – John pushed Max before he fell. In fact, the key to the interpretation of this reading's temporal ordering is that there is a causal relation between the events, thus enforcing the temporal order where pushing precedes falling. (Lambalgen, Baggio, & Hagoort, 2007) gives a more formal account of how this is done in terms of the event calculus, which we will quickly summarise.

First, the bare linguistic material must be updated in the discourse. In this case, there are two events, both in the past that must be fitted into the context. This is the first goal in processing the discourse. In this case, there is both an e_1 (`fall(m)`) and an e_2 (`push(j,m)`) that need to be added to the discourse model. Imagine them as the following goals:

- `update discourse with past event` $e_1 = $ `fall(m)` `and fit` e_1 `in context`

- `update discourse with past event` $e_2 = $ `push(j,m)` `and fit` e_2 `in context`

This puts both events in the past, which is necessary given the morphology, but says nothing about how the events are related to each other. The key then is the second half each goal, which requires that the events be fitted into the context. Suppose the listener has no more information than what is given in (1). How is the listener then supposed to interpret the prominent reading, where the temporal sequence of events is contrary to the linear order of presentation?

For this we must draw upon both general planning knowledge, knowledge of causality as well as stereotypes of particular events. We can start with an ontology that distinguishes between *events* and *processes*. Keeping on an intuitive level, both sentences in the above discourse can be said to be about events. For example, the *event* fall, would be said to include the point where Max's position has changed from somewhat upright to lying on the ground. There is also a process, *falling*, which corresponds to the motion and change in position of Max. In plain English, we can say that events are things that *occur* or *happen*, while process can be said to be *going on* for a particular amount of time. Processes can also be said to obey the *principle of inertia*, which should be taken for its Newtonian connotations. If a process is going on at time t and between t and a later t' nothing happens to change matters, then the process is still going on.

In general, there is a causal relation between events and processes. The first is that if a process is going on for a given period of time (and nothing intervenes to stop the process), a culminating event may occur. In this case we can distinguish between a *falling* process and the culminating event *fall*. The former would include the gradual change of angle, perhaps accompanied by curses, etc. The process, if left undisturbed, will lead to the subject flat on the ground. To

illustrate the difference between the process and the event, it is quite easy to picture the former with the latter never happening.

(2)   Max began falling. He quickly regained his balance and punched John in the face.

So far we are no closer to a link between the two events in the discourse; this is where the second causal relation between events and processes plays a role. Events can begin or end processes. In this case we have reasoned from a *fall* event happening to there being a prior process of *falling*. Moving further backward in time, the hearer can also reason that *falling* probably did not begin by itself, but was caused by something. In general, the process can be initiated by such things as slipping on a banana, losing one's balance, or even being pushed.

The latter possibility is the most germane, given the lexical material, and the hearer will then make the link between the process *falling* and *being pushed* and know that generally the latter temporally precedes the former. This is almost enough for the hearer to interpret the causal reading. Once the connection between the *push* event and the *falling* process is made, causality tells her that the *push* event must precede the process *falling*, which precedes the event *fall*.

But there is one caveat. From sentence (2), we have seen that the *falling* process does not necessarily lead to a *fall* event. Another event can intervene to delay or prevent it from happening. In fact, it is possible to have a *push* event, a *falling* process, and a *fall* event, but the causality not quite so straightforward. Take this convoluted example[2]

(3)   John pushed Max and he began falling. John then reached out an arm stopping him, but yet again changed his mind and let him go.

Even if the hearer does build her discourse model in such a way that the *push* event is temporally first and triggers a *falling* process, a final reasoning principle needs to be brought into play. This is called *closed world reasoning*, which says, 'assume all those propositions to be false which you have no reason to assume to be true.'[3] The hearer has a general causal connection between *falling* and *fall* as well as general knowledge of the principle of inertia. The hearer can then almost assume that once falling has started it will lead to *fall*, and that the *falling* was initiated by the temporally prior *push* event.

But, as sentence (3) shows, it is possible that the *falling* process was terminated prior to the *fall* event being culminated. This is where closed world reasoning comes to the rescue. Given only sentence (1), there is no reason for the hearer to infer that an extraneous event did occur.[4] The hearer then has enough in her armoury to interpret the mini-discourse in the non-linear, causal way.

---

[2]Perhaps imagining a good angel and bad angel both telling John what to do.
[3]Elegant formulation taken from (Lambalgen et al., 2007).
[4]In fact, it would take a fairly heavy construal to manage such a scenario with (2).

This sort of interpretation procedure is quite automatic most of the time. We do it multiple times a day without giving it a second thought. Nevertheless (as can be seen in the above), it is not so easy to describe, and successfully formalising such a thing is a no easy matter.[5] A simple two sentence discourse with its default interpretation is actually a rather complex entity, requiring the ability place events in time, integrate them in the proper way, applying planning, causality, and specific world knowledge that can connect the events. Now imagine the child in the situation who must both have the conceptual ability and conventional knowledge to perform the above operations, as well as connecting this to the way her specific language works. This is a daunting task indeed, but one that most of us manage to master. In the interlude following this chapter, the formalism of the event calculus is detailed, as well as how a formalism that takes these cognitive problems seriously can give a computational account for how they are done.

This chapter shall be an exploration into a variety of psycholinguistic experiments into how children both develop the conceptual apparatus necessary for tense and aspect comprehension and production as well as how these concepts can be mapped to the language they are acquiring. The reader is to be forewarned that they should expect no specific answer to language acquisition problem, but nevertheless will emerge with a bit of illumination about how complex the process is and the necessary ingredients involved.[6]

Before the language acquisition studies are examined, the preliminary sections can be seen as the foundation. In the above, it is clear that the comprehension of tense can be seen as having two primary goals: locating the event in time, and integrating with other events, present either in the discourse or the general context. This then raises the question as to what an event is – a question elegantly answered by the work of Zacks and Tversky. Again, notions of planning, goal and intention come into play as well. Finally, the reader is encouraged to note the close relation between the concepts and process involved in speaking of and understanding speech about events and the mechanisms of the event calculus, though more shall be said in the conclusion.

---

[5]A formal analogue of the above convoluted description will be seen before moving to the more formal chapters of the thesis.

[6]As mentioned in the *How to read* section of Chapter 1, the next two sections deal with the psychological counterpart of constructing event types. The sections that come after deal with acquisition of tense and aspect and are of less relevance to what comes in the latter chapters. However, should the reader persevere, the section on event on object individuation (Section 7) connects tense and aspect acquisition with event-type acquisition, and is a useful perspective for viewing the scenarios discussed formally in the Interlude. Finally, the reader should also be reminded of the caveat in the *How to read* section. This chapter is to give those of a semantics background a knowledge of the parallels that exist between the conceptual typology of Aktionsart and its psychological counterparts, as well as to provide a grounding for the formalism. It is in no way meant to be a work in the field of psycho-linguistics.

3.2 Preliminaries: Analogy between objects and events

(Zacks & Tversky, 2001) initially defines an event as "a segment of time at a given location that is conceived by an observer to have a beginning and an end." The process by which observers identify the boundaries and their relationship is called *event structure perception*. Of course, this definition does not conform to either naive usage of the term 'event' or various semantic definitions. First, as it is tied to perception, it can only apply to individual instances of event-types, i.e. event *tokens*. Second, both naive and semantic conceptions of events allow for both tokens and types with perhaps fuzzy spatial or temporal boundaries, temporal discontinuity, and time spans that would be impossible for an observer to observe. For example, to go back to the stock *building a house* example, one would not expect the preparation-phase building activity to be continuous over time, but stopped and started each weekday. Furthermore, a trip to local DIY store for equipment could reasonably be considered part of the preparatory activity and yet is not spatially connected to the building site. Also, the 100 Years war is certainly an event, but even if it were spatially and temporally contiguous, it went on for far too long for one observer to observe the relevant boundaries.

Nevertheless, there is something to be said for this definition as a starting point, and perhaps as a basis for developing the more higher level or somewhat abstract notions of event as seen above. Moreover, taking this as a starting point we can more easily see some of the important parallels between the perception of objects and the perception of events.

Treating events as somehow on a par with objects has a bit of a pedigree in both the philosophic and psychological literature. (Quine, 1996) proposed to go as far as simply treating events as objects, i.e. a bounded region of space time. That is, what are thought of as ordinary objects such as cars, rocks and trees are one type of bounded space-time region, and events another.[7] (Miller & Johnson-Laird, 1976) made a similar proposal by treating events as 'dynamic' objects and 'objects' (the naive world of rocks and trees) as concrete objects.

Objects are recognised by shape, texture, colour, etc. They also have boundaries in space, and one can (usually) quite easily identify what regions of space the object occupies and what regions it doesn't. Analogously events have boundaries in time, that is they have a beginning and an end. Like objects, they are also bounded in space. To use Zacks and Tversky's examples, a coffee cup (object) takes of a an amount of space of a particular shape. Picking up a coffee cup (event) happens in a particular location in space, that is also spatially bounded.[8]

---

[7]Again, this notion only captures event-tokens. It also side-steps many thorny analytic metaphysical problems with event identity if actions are brought into the picture. We shall quite happily also side-step such metaphysical considerations.

[8]Simplistically one could imagine a film of said event slow enough that each still still leaves

Continuing the analogy, we shall look now look at how both objects and (simple) events can be identified and categorised.

3.2.1   Taxonomies and Partonomies

There are two types of hierarchical structures that can be used to characterise objects, and by extension also events – taxonomic and partonomic structures. A taxonomy, well known from biological classification, is a kind-of structure. For example, a Smart Car is a kind of automobile, which in turn is a kind of vehicle. On the event level, a game of five-a-side football is a kind of "football", which in turn is a kind of "sport". Both objects and events can be referred to at these various levels of abstraction. A common feature of this analysis is to take there to be three levels of abstraction – subordinate, basic and superordinate. In the examples above, the basic levels would be automobile and football, the subordinate the more specific Smart car and five-a-side football, and superordinate the more abstract "vehicle" and "sport".

This division and categorisation may seem a bit trivial or obvious, but it does have some psychological impact to it. Rosch et al[9] have shown that the basic level is the psychologically privileged level in that it is the easiest to access in perception, action and language. It also conveys the most amount of information (compare "table" to either "coffee-table" or "furniture"), and members of this basic level category have more similarity (in shape for objects) to each other than members of the other two levels. Experiments done by (Rosch, 1978) show that name of objects is fastest at the basic level and that basic level concepts also tend to be the earliest concepts children learn.

Note also that a taxonomic hierarchy can help in inferences and reasoning, especially about intrinsic properties. For example, if one knows that a "salmon" is a member of the category "fish", then one can reasonably infer that a salmon has gills and swims in the water. In this case, this is true, but these type of category inferences are certainly defeasible. One need only look at the plethora of non-monotonic logics designed to defeat the inference that Tweety can fly if Tweety happens to be an atypical bird like a penguin or ostrich. We shall soon see as well on the event level that similar category related inferences can also aid in reasoning. In experiments carried out by Zacks, Tversky and Iyer (Zacks, Tversky, & Iyer, 2001) on event segmentation, there is an interaction between taxonomic knowledge and the ability to segment an unfamiliar event.[10]

---

a ghost image tracing out a path of the movement of the coffee cup.

[9]See (Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976) and (Rosch, 1978).

[10]Subjects with no musical training fared about as well as Stanford band members when asked to segment "putting a saxophone together" into component parts. Presumably, the ignorant students at least were familiar with the process of putting something together, and could make the proper inferences.

The second type of hierarchical organisation is a partonomy. Objects can be broken down into component parts. For example, a table has a top and four legs, a car has doors, windows, wheels, seats, etc. In cases like the car, the main parts just listed can be again broken down further into sub-parts, such as the seat, which has a back, headrest, bench, etc. Observe, that all the parts and subparts just listed have distinctive spatial configurations. In fact, shape plays a primary role in object recognition. Furthermore, parts when joined together give rise to discontinuities of contour, which may aid in object recognition. For example, it is possible to imagine two different objects consisting of the identical parts but put together in a different way (say a typical car and a rather odd configuration of car parts). These two objects would presumably be recognised (or not in the latter case) as different types of objects. In this case it would be a car, and a rather fine example of Art Brut, given the radically different contour discontinuations. Of course, further inspection would then yield the revelation that the art piece is indeed made out of car parts (drawing on the knowledge of the partonomic hierarchy of a car).

Just as a taxonomy can aid in inference, knowledge of partonymic hierarchies can help in reasoning from physical structure to function. If we see an object with legs, it seems reasonable to infer that the object can stand whether the object is inanimate such as a chair or table, or animate like a zebra or human.

Breaking events down into component parts was already seen in the previous chapter in the discussion of event structure. An accomplishment could be seen as having parts such as a preparatory activity, an achievement, and resultant state, all of which could also be broken down into component parts in a recursive manner. Similarly, on the psychological side, events are broken down into hierarchical, component subparts. For example, using a script level of story understanding, the cognitively privileged partonomic level can be considered scenes in a script. A simple illustration would be the script for dining in a restaurant. The scenes would then be such things as "entering" "ordering", "eating", "paying the bill", etc. The subordinate level would consist of the various sub-events of the scenes (i.e. "ordering" could be a series of conversation exchanges punctuated with some menu-pointing).

Observe that the boundaries of the parts can also be identified with the locations of maximal perceptual change, in a way similar to parts of an object being identified by discontinuous spatial contours. Similarly, the sub-events would also be characterised by discrete perpetual changes. Keeping to the restaurant theme, the "entering" scene could be subdivided into opening the door and stepping in, being greeted by a host, then walking to a table, and finally sitting down. Should these sub-events run smoothly, it would be rather strange to break them down further, like the walking to the table consisting of two further sub-events. This is conceivable, however, if the walking is interrupted by a streaker running across the restaurant. Again, however, there would be a rather dramatic perceptual change.

This kind of characterisation illuminates a subtle interaction among a number of factors. The first is the cognitive-perceptual system which breaks down the stream of perception into manageable bits. With this alone, there doesn't seem to be enough to determine if a given action is on the subordinate or scene level – as both would be based around this 'discretising' of perception. Characterising what is a sub-event and what is an event can be seen in broad terms as 'cultural' and based around knowledge of intentions (again another supposed cognitive mechanism), causes, social convention and world knowledge. For example, ordering and paying the bill could appear quite similar (purely perceptually) to an observer with no acquaintance with ordinary human activity. The difference lies in the goals and the ordering of the scenes within the entire script, which is largely cultural knowledge. It is then no surprise that (Bower, Black, & Turner, 1979) experiments show that within 'relatively homogeneous' samples of people share agreement about the typical parts of everyday activities.

3.2.2   Event Boundaries over time

Ultimately, whether people will primarily characterise an event on the basis of perceptual features and discontinuities or in a more cultural or goal based manner is largely based on time scale. Taking the example of a handshake from (Zacks & Tversky, 2001) we can see how this fairly simple event can both be broken down into component parts and be part of ever larger, more abstract or temporally extended events. The smallest temporal scale that of a psychologically reified event is of a very short time scale, and is primarily defined in terms of simple physical changes. In this case, the events could be the raising of the hand, grasping of the other person's, etc. On a slightly longer timescale (10 to 30 seconds), events can be defined in terms of a simple, intentional action, in this case, the handshake as described above. Moving further to a time scale of minutes to hours, there are events characterised by goals and plans and socially conventional forms of activity. Here the handshake may now be a subpart of a treaty signing. Finally, on an even larger scale there a events that are best characterised thematically – the treaty-signing as a sub-event of a peace process.

Notice that by the time the top level of the time scale is reached, the link to simple, physical changes is gone. It is best described by (Zacks & Tversky, 2001):

> In general, it seems that as the time scale increases, events become less physically characterised and more defined by the *goals, plans, intentions and traits* of their participants. (p.7) (emphasis added)

Of course, that is not to say that physical events have nothing to do with events characterised at the thematic level. World War II for example, in part, consists of innumerable events that could be characterised at the various levels of description. But the point is that even if it were possible to describe all of the

sub-events of the war on the basis of simple physical changes,[11] it would not be a proper or satisfactory answer to the question of what World War II was. At a minimum one would have to name the parties involved and their respective plans and goals. Even a simple intentional act such as a handshake, while made up of the smallest reified events such as described above cannot be explained merely by these characterising events. The purpose (in a social context) of the handshake must also be brought into the explanation. Intentions and goals, therefore must be taken to be also intrinsic in people's perception and conceptualisation of events.[12]

3.3 Intentions and Goals

In Chapter 2, the importance of the notion of goal to account for certain event-types was shown. It is no less important in the psychological realm. This is especially true if we take 'psychological' events to be a (temporal) partonomy as described above. Goals can be recursively seen in terms of sub-goals, generating a hierarchical partonomic structure. Using the example in the last section we can see the now famous handshake as a sub-goal in the signing of the peace treaty, and the peace treaty as a sub-goal in achieving the higher level goal of a successful peace process.

Again, a nice feature of these sort of hierarchies is that the allow default inferences based on knowledge of event scripts or schemata. Zacks and Tversky give a cute example of the default inference of an observer in a bank who witnesses someone enter carrying a gun and wearing a mask. The observer will most likely infer that the intruder has a goal to rob the bank, and yet a higher goal of acquiring money. Moreover, knowing the goal allows for both the prediction of future behaviour and inference from incomplete information. For example, the observer can predict the 'robber' will make a move for the safe or a bank window, knowing the robber's goal of money acquisition. She will also be able to infer that if the robber is exiting the bank with a bulging sack, this sack contains money, gold or some other valuable item, and probably not a pile of yesterdays newspapers. Of course, like the inferences from the related hierarchies, it is defeasible. Our robber may only be a performance artist, or perhaps an actor hired to test the bank's security procedures.

The importance of goals and intentions in both the linguistic and psychological account of event structure is no accident. Some authors (e.g. (Tomasello, 2003)) take the ability of humans to understand the intentions and goals of others to

---

[11]How this could be done without reference to borders and other conventional concepts makes this attempt rather doubtful.

[12]In this I am ignoring that fact that any voluntary bodily movement can be seen as an intentional act. The difference seems to be that one can explain, say, what 'grasping a hand' is in purely physical terms and no mention of any sort of higher level social intention or goal (of course there is the goal of actually grasping the other hand). Answering the question of what a handshake is requires both the physical description and social context to be satisfactory.

be the primary human evolutionary leap from our fellow primates. This ability allows for all kinds of cultural learning, including the acquisition of language.

From about 9-12 months of age infants can already understand the communicative intentions of others at least to some extent. (Meltzoff, 1995) shows that 18 month-olds can already distinguish between manner and goal in human activity. When shown adults attempting to achieve a goal (both successfully and unsuccessfully) the children will only produce the action that lead to the achievement of the goal. When shown an inanimate object 'performing' an action, children produced the action whether or not it lead to the successful achievement of the goal.

(Tomasello, 1999) cites a fascinating experiment in (Gergely & Csibra, 1993) that is a wonderful example of intention and plan reading and its impact on imitation by 14 month old infants. The infants were shown an adult bending over to push a button for a lamp with their head rather than in the ordinary button-pushing way. In one condition, the adult acted as if the room was cold and put on a shawl that was held by both hands.[13] When the infants were asked to turn the light on in this condition, they turned it on in the ordinary way (with their hands). In a different condition, infants saw the adult again turn the light on with her head, but this time there was nothing impeding the use of her hands. In this case, the infants imitated the adult's odd way of turning on the light, presumably, because they thought that this time there was a good reason to do so that they wanted to find out about. In other words, in the second condition, the adult must have had a good reason for acting in such a way, and it was too tempting to achieve only the goal, without also imitating the means.

Furthermore, experimental evidence from (Wagner, 2001) shows that when under five years old, children are primarily using their knowledge of intentional behaviour to compute the completion possibilities of telic predicates. This will be described in more detail when looking at various theories of grammatical tense and aspect acquisition in the context of the aspect-before-tense hypothesis.

Before moving on to the experimental evidence that humans really do encode ongoing activity (and in memory) in the hierarchies proposed so far, it is useful to sum up. First, there is the perceptual mechanism that segments both objects and events (also seen as 'objects' in this view) in terms of discrete spatial (and temporal contours). This bottom level can be used as the basis for the development and acquisition of event schemata in terms of goal, subgoal hierarchies. However, it should be emphasised that these are to be seen as general cognitive mechanisms that are neither unique to language or event perception and structuring. If these two mechanisms have evolved by natural selection, the specifics have evolved by cultural evolution.

> It seems likely that over evolutionary time some goals have re-
> mained salient, particularly the general goals that can be characterised

---

[13]Making their arms temporarily unusable.

as drives: to eat, to procreate, to avoid harm. However, for much interesting behaviour, the relevant goals and plans are likely to be almost entirely culturally transmitted. ((Zacks & Tversky, 2001), p. 15)

So, the only universality lies in the *way* we structure and represent events, not any event schemata in particular. These will vary over time, cultures, different people in the same culture, and even in the same individual at differing times in their life. With that said, it is time to look at experimental evidence supporting this notion of event segmentation and its possible ties to language.

3.4 Event Segmentation and the Hierarchical Bias Hypothesis

Thus far we have seen certain notions such as taxonomic and partonomic hierarchies and event scripts or schemata as possible ways in which events are perceived and organised psychologically. And while the notions do seem to have some intuitive plausibility and would be a clever way to have a computer 'understand' narratives, we have yet to see any evidence of their psychological plausibility. This section will remedy this situation. We will see numerous results from experiments that indicate that, indeed, there is evidence that events are both perceived and stored in memory in the aforementioned hierarchical structures. However, questions remain as to how far this model can be taken, as well as the relation to the proposals that events are organised on the semantic level in a strikingly similar way. This section will culminate with experiments recently done by Zacks, Tverksy and Iyer[14] on event segmentation by Stanford undergraduates, but first we shall see some results from previous experiments regarding events and narrative interpretation and memory.

3.4.1 Plans, goals and subgoals

On the narrative comprehension and exposition side, there have been numerous experiments that have shown that people will organise such narratives according to a hierarchical goal, sub-goal structure. This tendency does take time (in development) to come about, however. (Trabasso & Stein, 1994) looked at narratives produced from a picture book about a child and his dog searching for a wayward frog. The youngest subjects (from about three to five years old) kept their 'narratives' more or less restricted to descriptions of characters and relations in the individual pictures. But, as the children approached nine, there was increasing use of a goal-plan structure in the recounting of the picture book – the top goal

---

[14](Zacks et al., 2001).

being the child's desire to get his frog back, with the various sub-goals being the various attempts, searches, etc. By nine years old, children made as much use of a goal/plan structure as did adults.

But, recall that children as young as one year are able to understand goals and intentions as well as differentiate manner and goal. So perhaps, there is gap between the understanding of basic goal, sub-goal relationships and the ability to construction a conventional narrative.[15] Nevertheless, children as young as two have been shown to have the ability to manipulate and understand somewhat complex goal, sub-goal structures.

(Travis, 1997) did experiments that do to show that such hierarchical representations of events are used to guide perception and memory. The children were shown brief sequences that contained two actions that would lead to a goal action. To make matters more interesting, the two relevant precursor actions were mixed with irrelevant actions. An sample sequence would be: "(1) hang up a gong, (2) place feathers in a container, (3) make a clapper, (4) remove a plug from a bellows"[16] with the goal action being (5) banging the gong. When asked to replicate the actions, the toddlers most often rearranged the actions so that the relevant sequence (in this case, 1, 3, and 5) were performed sequentially. Also, more of the goal-relevant actions were produced than goal-irrelevant actions.

Recall the link between a goal hierarchy and the notion of a partonomic event structure. Both break down down into a iterated structure, with the schemata encoding the relevant information about goals and sub-goals, and these breaks aligning with the simple partonomic structure. For example, there is the toy example of the script for eating in a restaurant. Broken down into scenes and their subparts we obtain a simple partonomic hierarchy. But, also recall the names of the various scenes – "entering", "ordering", "paying the bill". These scenes can all be seen as goals or subgoals a successful restaurant event. This is precisely what is meant by the schemata aligning the goal information with the part-whole structure. While the initial stages of such schemata are evident in infants and toddlers, as can be seen by the Trabasso and Stein experiments, full-blown event representations take some time to develop. Now the claim is that these schemata influence both action and planning in that the schemata in that the schemata play a role in how people think about their own actions in the past. And indeed there is evidence that both memory of perceived events on video, as well as autobiographical memory is indeed shaped by event schemata.

---

[15]There is a good possibility that the youngest children don't quite understand the task at hand, and are merely content to describe the individual pictures, not quite have grasped the concept of a narrative. Parent/child interaction at this age may play a significant role, as well.

[16](Zacks & Tversky, 2001), p. 16.

3.4.2   Event Schemata in Memory

Sticking with narrative for the time being, experiments by (Bower et al., 1979) indicate that encoding events schematically can influence reconstruction in memory. In one experiment, subjects were give a script-based story. Afterwards, there were given a memory recognition test, including descriptions of events that were not in the story read. If these events were plausible parts of the script in the story, there were high false recognition rates – that is, the participants thought they had read a description of a particular event if it was plausibly a part of the script. A second experiment was to present the event parts of a particular script in an a typical order and then have the subjects recall what happened. Unsurprisingly, participants tended to remember the events in the conventional order. This sort of false reconstruction does seem to be hierarchical, though. In a similar study by (Abbot, Black, & Smith, 1985), subjects were more likely to infer false actions when presented with a description of a scene, than to falsely name a scene based on an inference from give actions.

Moving closer to actual on-line perception, memory of videos of human activity also show similar effects of a hierarchical structure. Recollections for actions relevant to a particular event schema are more likely than those of irrelevant distractor actions. (Lichtenstein & Brewer, 1980) show reordering of event parts from atypical to typical order in a similar way to the (Bower et al., 1979) results with a recollection of narrative. Confirming the findings of Bower et al., subjects also tended to falsely remember viewing actions that were not shown in the videotape, but were nevertheless implied by the event schema represented.

Finally, even autobiographical memory shows evidence of hierarchical patterns. (Anderson & Conway, 1993) have shown that when a certain episode has been activated in memory, it subparts become more available for recollection. Also, over time, specific information deteriorates, and memory of autobiographical events takes on an increasingly schematic nature. (Barsalou, 1988).

So, we now have compiled a good body of evidence that, at least in recall of various mediums, there is a significant bias of hierarchical structures in relation to events. This is something that develops in children over time, starting with an understanding and ability to manipulate simple goal, precursor action structures. Over a period of a few years, full blown event schemata develop which combine goal and causal information to align with the boundaries of a conventional part-whole structure.[17]

But recall the first stab at defining an event psychologically – "a segment of time at a given location that is conceived by an observer to have a beginning and an end." The most basic notion of event, then, involves on-line perception, thus asking the question as to whether hierarchical patterning might not also affect the encoding of events while they are being perceived. Perhaps we do not encode

---

[17]Though, of course, the partonomic boundaries at higher levels of granularity (time scale) will be increasingly defined in terms of goals and causes.

events hierarchically as they are observed, but because of conventional ways of telling stories and communicative needs, a sort of hierarchy is activated at the point of the recounting. For enlightenment on this questions we will look at the experiments of Zacks, Tversky and Iyer designed to test the hierarchical bias hypothesis.

3.4.3   On-line event segmentation and the Hierarchical Bias Hypothesis

The Hierarchical Bias Hypothesis is simply the notion that people will spontaneously be disposed to actively encode ongoing activity in terms of a hierarchical part structure. Testing this is the primary goal of experiments first by (Newtson, 1973), and modified and extended by (Zacks et al., 2001). This is done by having participants mark out "naturally occurring" event boundaries while watching videotape of various activities. The extensions, however, are of particular interest here, because they involve linguistic descriptions (by the subjects) as well. With this they could test the extent to which linguistic represents may actually activate perceptual representations. That is, in order to produce a adequate description one may need to cross temporal granularity levels (say, describing actions in a larger context to make sense of them), as well as activate representations of goals and plans of the actor doing the activity. In other words, how well do the linguistic descriptions reflect a hierarchical structure? Note that this assumes that language and the cognitive representation of events are closely linked, and does fit in rather nicely with the semantic event structure given in Chapter 2.

The basic format of Newtson was to have subjects view a videotape and press a key to mark when they observe "natural and meaningful" unit boundaries. One group of participants is told to mark the largest natural and meaningful units, the other group, the smallest units. The initial findings were that points marked by the large unit condition group also tended to be boundaries for the small unit group. Also, points not marked by the small unit group also tended not to be marked by the large unit group. To make sense of this, let us once again return to the restaurant script. The large unit boundaries would correspond to the boundaries between the various scenes, while the small unit boundaries would correspond to the next level down on the hierarchy – the various events that make up a given scene. The alignment would then be where sub-events and scenes meet, i.e. the inceptive or terminative events for a scene. This seems to be exactly the hierarchical structure that was being looked for. At least, if there is a hierarchical structure in the encoding of event perception, one would certainly expect such strong alignment effects.

Unfortunately, the initial experiments engendered a good deal of controversy, with arguments that these segmentation tasks showed no such bias effect. The

objections were both on the specifics of the tasks as well as more theoretical grounds. For example, Newtson himself[18] holds that even if the hierarchical pattern found in the segmentation tasks is acceptable, there is no need to conclude that there is any cognitive representation of events organised in such a hierarchy. In other words, the alignment pattern is not enough, one also needs to investigate the interactions with memory, language and prior experience.

Zacks, Tversky and Iyer conducted five different experiments, designed not only to test the online segmentation, but also the needed interactions with language, memory, familiarity and communication.

Main Experiment

The first experiment involved videotapes of four different, single actor intentional activities (each divided into a script of twelve steps): assembling a saxophone, fertilising houseplants, doing the dishes and making the bed. The first two were ranked unfamiliar (by a survey process) and the latter two, familiar. Like Newtson's experiments, subjects were asked to tap a keyboard when they encountered a smallest or largest natural unit boundary (depending on which group they were in).[19] As an added twist, 32 of the forty participants were also equipped with a tape recorder for which they should utter a short description of what was happening when they tapped out a breakpoint.

The analysis of the data indeed shows the presence of an alignment effect. That is, "unit boundaries under the coarse and fine coding conditions were in better alignment than would be predicted by chance." (p. 34) Moreover, there was some interaction with familiarity. The breakpoints for the coarse condition (i.e. largest unit boundaries) were slightly more aligned with the breakpoints for the fine condition (smallest unit boundaries) on the familiar events.[20] Given that individual scripts must be learned, an interaction with familiarity is to be expected.

Finally, there were the effects of verbally describing the activity while doing the segmentation. (Zacks et al., 2001) note that there are two possibilities of what the effect might be. One possibility could be that there is interference between having to make a description and having to mark out the segments by tapping. This would lead to a weakening of the alignment effect between coarse and fine boundaries. On the other hand, if there is some validity to the hierarchical bias hypothesis (i.e. people 'naturally' encode activity in terms of schemata) *and* if such schemata are somehow constituted by linguistic or semantic repre-

---

[18](Newtson, 1993).

[19]Actually, subjects (after a distractor exercise) were asked to encode at the opposite level unit boundaries in a second run through.

[20]The difference is a matter of milliseconds, however. And their analysis, admittedly, had the disadvantage of not being able to rule out extraneous factors in the different activities that could also account for the difference.

sentations, then describing the activity may actually strengthen the alignment effect by increasing the influence of the hierarchical representations. Indeed this is exactly what happened – describing the activity did correlate with an even more pronounced alignment effect, and the same hierarchical structure present in the non-verbal segmentation, also shows itself in the verbal activity descriptions.

There were also some interesting results regarding the use of language. Ratings were given to the various descriptions including verb generality and goal-directedness, object generality, ellipsis, pronominalization, etc. Furthermore, the fine unit descriptions were divided into two subclasses – *boundary units* and *internal units*. The boundary units are the fine unit boundaries that also corresponded to coarse unit boundaries. Under the hypothesis that linguistic description of activity is also subject to the hierarchical bias effect one may expect some correlation in description between the boundary units on the one hand, and the coarse unit boundaries on the other. And, based on the rating system sketched above, when calculating the difference between the coarse and fine boundary descriptions, the boundary unit descriptions were just as close to the coarse unit boundary descriptions as the were to the internal units. This is taken to be evidence that even when describing events at the lower level of description, a higher level (coarser temporal grain) is also being accessed.

To give a simple example, the boundary units that corresponded to the coarse unit boundary of "putting on the blanket" in the making-the-bed activity were described as "spreading out the blanket" and "tucking in the blanket". The internal units that reference the blanket do so not with the full object noun, but rather a pronoun – e.g. "straightening it". Note that "putting on" is more general than "straighten" or indeed than the verbs for the boundary units "spread" and "tuck". So, in this example, verbal usage in the boundary unit descriptions do correlate with their fellow fine unit descriptions. But on the other hand, they are closer to the coarse unit descriptions in that they use full nouns rather than pronouns.[21]

In general, the coarse-unit descriptions specified the objects much more, and contrastly, the fine unit descriptions put the specificity on the verbs, but there seemed to be no effect from verbal aspect (i.e. simple or progressive usage). Use of polysemous verbs occurred more often in coarse unit descriptions and boundary unit description (though less here). The opposite scale corresponds to polysemous nouns, which occurred most with internal unit descriptions and least with coarse unit descriptions, with boundary unit descriptions falling somewhere in between.

---

[21]Of course, one could object that once one starts the description of the "putting on the blanket" scene in the fine description condition, the full noun would naturally only be used the first time (in common linguistic usage), and the following internal unit descriptions would naturally use a resumptive pronoun. The use of the full noun at the end might seem more powerful evidence, but for the fact the description before the end-boundary referred to "lifting the bed". The speaker then wishing to refer to the blanket again would no longer have the pronoun accessible.

However, there were some common linguistic features that cut across all classes of unit description. In both cases the VP's tended to refer to either accomplishments or achievements (i.e. goal related), rather than simple, physically bounded event descriptions. That is, when asked to segment and describe the events in terms of the "smallest" natural units, the participants could have gone as far down as such descriptions as "raising the arms" or "bending the waist". Instead, even the internal boundary descriptions were such things as "putting on top end", "picks up pot", or "opening a drawer".

The conclusions drawn by the experimenters are that having to actively describe each segment invokes hierarchical knowledge structures as well as "greater awareness of causal, functional, and intentional relations.", and that there is an interaction between raw perceptual differences in activity as well as knowledge of event structure. A question not addressed by this experiment would be whether in the non-linguistic condition, do subjects exhibit even finer grained boundaries in the "smallest" condition that would correspond to the basic physical events? Or does having language already condition the encoding to look a bit higher up on the hierarchy? (Zacks & Tversky, 2001) do hint at an answer themselves. "This in turn suggests that using language, *and perhaps language itself*, biases away from raw perceptual statements and toward causal and intentional ones" (p.41).

Interactions with familiarity

The second and third experiments were designed to take a look in greater detail at the interaction of familiarity in event segmentation. Given that on the cognitive model described above, even if the hierarchical structuring is universal, any specific schema would have to be learnt. Therefore, one would expect that familiarity with a particular event sequence would strengthen the alignment effect.

In Experiments 2 and 3 familiarity was made even more manifest. Taking "assembling a saxophone" as an unfamiliar activity, in Experiment 2, the subjects were first taught how to assemble a saxophone and then made to do the event segmentation task. In Experiment 3, there was no special master-class, but rather two groups were compared – members of the Stanford Marching Band and saxophone novices. So, the expected results would be an increased alignment effect in Experiment 2 as opposed to Experiment 1 (the recent teaching adding to knowledge of the script), and the Stanford Band members have a greater alignment effect than the novices.

However, this did not turn out to be the case in either experiment. Now, this could be interpreted to evidence against the Hierarchical Bias Hypothesis. But, on the other hand, one also could interpret as the difficulty of doing experiments. Recall the discussion earlier on Taxonomic Hierarchies and the ability to make inferences from them. In this case it just may be that even to novices the event of assembling a saxophone is really not so unfamiliar. All the Stanford students can

be assumed to have some knowledge how objects are in general put together from component parts. This knowledge, which would be a level up on a taxonomic hierarchy from "assembling a saxophone" could then be applied to the event segmentation task at hand, thus nullifying any effect from the participant being relatively unfamiliar with the specific activity.

Structure in Memory and Communication

The final two experiments brought both memory and the comprehension of event descriptions into the frame. The memory experiment involved the same video-taped activities, but instead of online perceptual (and linguistic) segmentation, the subjects were required to write down the event segments immediately after watching the activity being performed. Again they were asked to write the events down according to either the largest or smallest possible "natural" and "meaningful" units. Note that a rather important difference from the first experiment is that temporal encoding no longer plays a role. Thus, there is no way to divide the fine unit descriptions into boundary and internal units, in that there is no way to temporally align the coarse unit descriptions with the fine ones.

Nevertheless, the descriptions given from memory were quite similar to those produced in the on-line perceptual experiments, and the authors are able to come to the conclusion that the general pattern of the descriptions from memory also supports the hierarchical bias hypothesis.[22] However, there were some linguistic descriptive differences that could be the result of the experimental setup, i.e. writing from memory as opposed to an impromptu utterance. Furthermore, there was one marked difference in that the verbal descriptions in the coarse-level descriptions were more specific than their fine counterparts, the exact opposite of Experiment 1. This may be as a result of the different linguistic mediums or perhaps that when divorced from the perception of immediate physical activity, the schema take up more of a role and the description focusses more related to the goals and plans of the agent, and the more specific verbs carry the specific goal of the action along with it.

Finally, the last experiment involved giving subjects transcripts from the fine unit conditions of the first experiment. The participants were then asked to divide the individual items into larger, continuous groups. The prediction was that even in the comprehension of event communication the hierarchical bias hypothesis would play a role, and the groups identified would correspond at least roughly to the course-unit segmentation from the first experiment. In fact, this is exactly what happened – the subjects of this final experiment created boundaries very close to those of the participants in the first experiment's coarse coding condition. Reasons for this may be that the participants do have the requisite background knowledge of the tasks involved, and that even these sparse descrip-

---

[22]Recall also from previous sections the experiments dealing with memory and event structure.

tions are structured in such a way that the syntactic and semantic cues in the terse descriptions are enough to guide the reader to similar breakpoints. Remember that the analysis of linguistic descriptions from Experiment 1 showed that the boundary units were linguistically described in a way that was half-course and half-fine. Again, the hierarchical bias may be used by readers to decode texts as well, with the reader assuming that the writer has a similar, hierarchically structured representation of events, and this guides interpretation.

3.5 Acquisition of Tense, Aspect and Aktionsart

3.5.1 Introduction

In the previous section we looked at experimental evidence regarding events in perception, memory and communication. While language, especially event reporting, played a significant role the focus was more on cognitive mechanisms that play a role in the perception of events. These mechanisms included a perceptual gestalt that leads to events being divided in a way analogical to objects based on changes or in spatio-temporal experience, allowing events to be divided into parts just as objects are. Secondly, a goal, sub-goal or planning mechanism is theorised to be used to generate a hierarchy on the part-whole structure. The connection with language then comes in when it is realised that this hierarchical part-whole structure is remarkably similar to the semantic structure given in the previous chapter. Ultimately, the claim is that language itself is effected by the above cognitive mechanisms in regard to the development of the semantic event structure given in Chapter 2.[23]

However, nothing in the event segmentation experiments deals with the way various linguistic systems are acquired. We have already seen some mention of the development in children of intention recognition, and goal-plan understanding. But what we yet to see is how this is put into action – i.e. how do children acquire the correct use of tense, aspect and aktionsart constructions? This section will attempt to remedy this. We will examine experiments by Weist, Wagner and van der Feest and van Hout[24] designed to test various hypotheses about how children from a variety of linguistic backgrounds understand and acquire tense and aspect morphology. While all addressing the same basic problem, the theoretical assumptions and methodology to test them do vary quite a bit.

At first glance, one may wonder exactly what tense morphology has to do with either the experimental results on event segmentation, the development of intention understanding by children, or the semantics of events. But, claims have been made in both the generative and functional camps that aktionsart and aspect plays a significant role in children's understanding of tense morphology.

---

[23]A similar claim can also be made for tense and aspect as well.

[24](Wagner, 2001), (Wagner, 2002) and (Feest & Hout, 2002).

This has generally been know as either the aspect (or lexical aspect) before tense hypothesis. But before examining the experiments in detail, some background regarding the acquisition of tense and aspectual systems is needed.

Articles by (Behrens, 2001) and (Wagner, 2001) cite numerous studies that children's early use of tense and aspect morphology, cross-linguistically, differs in patterning compared to that of adults or older children. Wagner gives the example in English that children are more like to use simple past tense for telic events and the imperfective *ing* form for atelic events. For instance, young English speaking children are much likely to say things like *broke* and *made* and *playing* and *riding* than *breaking* or *played*. Obviously, within a few years, these sorts of restrictions are relaxed, and verbs with a telic end-point are used with the progressive, and atelic verbs are used in the simple form, indicating perfectivity.[25] One way of conceiving this is that, for English, children are using simple past tense morphology to mean 'this is a telic predicate', and *ing* to mean, 'this is an atelic predicate', with the morphemes at this point in the child's development having no tense or aspectual[26] meaning whatsoever. Alternatively, as *ing* is purely a grammatical aspect marker,[27] and *-ed* carries tense information as well, it could be that children are ignoring tense and using the morphology to encode aspect only.

For a summary of possible explanations for the above phenomenon, including semantic predisposition and language specific theories, and the reader is directed to (Behrens, 2001) and (Wagner, 2001). Unable to cover everything here, we will concentrate on three groups of experiments that concern this issue. The first deals with a nativist approach, where an innate learning principle along with features of universal grammar are used to account for the trajectory and timing of the development. The second set of experiments are those of Weist, whose scope is broader than merely the above patterning, but presents a developmental account of children's development of temporal and aspectual concepts and how these are utilised to learn the tense/aspect specifics of their language. Finally, experiments by Wagner are discussed, the first of which test the Aspect First hypothesis.[28] Results from the first experiments, however, caused her to broaden the factors involved in morphology acquisition, including raising interesting questions relating to the differences of how children and adults conceive events.

---

[25]This example is for English, but (Wagner, 2001) cites studies that show this same kind of distribution of verbal morphology according to aspectual type in languages such as Italian, French, Japanese, German, Turkish and numerous others.

[26]i.e. lexical aspect, but not grammatical aspect.

[27]With the copula *is* or *was*, etc. carrying the tense information.

[28]The hypothesis mentioned above that tense morphology is initially used to encode grammatical aspect.

3.5.2 A Generative Approach to Tense Morphology Acquisition

Our first look at children's initial bias to use perfective marking (e.g. in English, *-ed*) with telic verbs and imperfective marking *-ing* with atelic verbs falls in the generative tradition. Using a variant of a learning principle detailed in (Olsen & Weinberg, 1999), (Feest & Hout, 2002) use this principle for experiments with Dutch children.

This is presented as the Syntactic Subset Principle. The crucial assumption here is that the mechanisms underlying child's grammar and adult grammar are essentially the same (i.e. Universal Grammar). The difference, then, lies in the exact settings of the innate grammar – that is, the child is scanning their linguistic environment to get the settings of their grammar mechanism to match that of their target language. Of course there are differences in a four year child's production and understanding of, e.g., tense and aspect. The difference arises because the child begins 'preset' with the most restrictive grammar possible and gradually relax their restrictions due to environmental input. Olsen and Weinberg write that this is a superior learning strategy as with the opposite – selecting the least restrictive hypothesis – the child "would overgenerate, producing examples that were not restricted by the rules of the language, and would need correction (negative evidence) to retreat from this overgeneralisation." (p. 530)[29]

For example, in Mandarin Chinese, grammatical aspect marking is restricted to certain lexical aspect classes, specifically, the imperfective only can be used for durative and dynamic predicates, while in Korean, perfective marking is restricted to telic predicates. These possibilities for marking are both far more restrictive than for a language such as English, where (with some restrictions – see Chapter 2), any aspectual marking is available for any type of predicate. Thus, the child starts out with the initial hypothesis to apply the imperfective to atelic predicates, and to apply the perfective only to telic predicates. If the child happens to be acquiring a language that has a broader aspectual marking pattern, the child will hear examples that violate the initial hypothesis and will expand their grammar accordingly. For example, an English speaking child will initially restrict the use of *-ed* (or irregular past) to telic predicates only, and then, upon hearing adult use of *-ed* with atelic verbs, will open up their hypothesis to allow for marking any kind of predicate with *-ed*.

---

[29]The literature is full of papers documenting overgeneralisation in children (and eventual retreat). This may be a less efficient learning strategy, but it certainly to happens (Chapter 5 of this thesis refers to an article by (Tomasello & Brooks, 1999) which looks at overgeneralisation of lexical causatives.) While correction certainly has no effect until the basics of the language have been learnt (i.e. when older children or adults are subject to social pressures for a 'correct' grammar in school), certain overgeneralisations disappear nonetheless.

Aspectual Class	Telic	Dynamic	Durative	Examples
State			+	know, have
Activity		+	+	march, paint
Accomplishment	+	+	+	destroy
Achievement	+	+		notice, win
Semelfactive		+		jump, tap

Table 3.1: Lexical Aspect

Subset principle for lexical aspect

(Olsen & Weinberg, 1999) proposes that in the early stages of the child's development, it is features of lexical aspect[30] that determine the distribution of grammatical aspect morphemes. Contrary to the view of event-types developing over time and influenced by perceptual, cognitive and social factors, both lexical and grammatical aspect categories are part of the innate Universal Grammar. The most restrictive application of grammatical aspect to lexical aspect (e.g. only -*ed* to telic, etc.) is justified by them as some languages do actually exhibit such a contrast. The relaxation of the restriction only occurs in language environments (such as English) where the patterning is not so restrictive.

However, while empirically there is a merit in the observation that children do initially use such a patterning, there are some problems in the formulation of lexical aspect, and more general questions as to what 'restrictiveness' is.[31] This is perhaps easiest to see if we look at their formulation of lexical aspect. It consists of features, *telic*, *dynamic* and *durative*. They combine to give the five basic event-types as can be seen on Table 1 (Olsen & Weinberg, 1999, p.531). The constraint then can be formulated to say that children restrict their use of *ed* to those categories marked +telic, and *ing* to those marked as dynamic and durative. This means that accomplishments allow both marking possibilities,[32] so the main issue is testing whether -*ed* is used with activities, states and semelfactives. The use of *ing* is always allowed as there are no verbs in their classification that are not marked for one of *dynamic* and *durative*.

An interesting case is the status of the semelfactive. The hypothesis is that children will first use, e.g. *coughing* only, and then relax their restrictions to use *coughed*. I suppose this is a possible developmental pattern for English. However,

---

[30]What we have been calling event-types or Aktionsarten.

[31]The following critiques are specifically aimed at the definitions, and empirical adequacy of their formulation of the subset principle, especially in regard to lexical aspect. As far as the feasibility of these sort of language specific learning principles, I have very little sensible to say. A question does occur, though. While the notion of innate semantic features may be appealing, is not the general process of matching particular verbs of their language to feature configurations still just as complicated with no such innate features and principles?

[32]Which they don't seem to mention, and when counting tokens multiple times, e.g. any scoring they give to accomplishments will never show any circumscription.

Chapter 2 notes that in Russian, certain semelfactives are only possible in the perfective. This would then predict that Russian children would first use the ungrammatical imperfective with these verbs and then relax their restrictions and acquire the proper marking. With no negative evidence allowed, it is a mystery as to how the children then stop their ungrammatical imperfective utterances.[33]

Relatedly, I do find the notion of restrictiveness a bit confusing. The justification for 'most restrictive' seems to derive from a notion of universal grammar, meaning that if individual languages exist with a more restrictive grammar, then all children start with this grammar. As the subset principle is not only applicable to aspect marking, but much more general, I do not know of what to make of a languages such as those of the Bantu family, which (Comrie, 1985b) notes have multiple past tenses, e.g. for *yesterday*, remote past and some time in between. This sort of tense system seems more restrictive, and would imply that English speaking children start out with these same tense categories and then (via input) realise that *-ed* applies to any point in the past.

Finally, Olsen and Weinberg analyse data from the CHILDES database to test their theory. There are token counts,[34] divided into developmental age stages, divided by the existence or lack thereof of each of the three features. Thus, for verbs with more than one positive feature, there must be double counting. In any case, the aggregated data does show a bias in favour of the subset hypothesis, but it is anything but categorical. For example, the first stage (presumably the youngest children) does have many more *-ed*'s for telics as compared to atelics but the raw token count is 537 to 159. Is this evidence of even the youngest children expanding their restrictions or the interaction of other factors? On a non-language specific account of learning, the lack of categoricity does not present a problem, but why are even the youngest children marking atelics with *-ed* if they are so predisposed not to?

Anticipating the objections to using aggregated data, individuals results for six children are given (but not their ages). Only one of the children shows categoricity effects (1 out of thirty uses of *-ed* used with atelics), while others go up to 30 per cent of uses of *-ed* with atelics. Interestingly, the child that conformed most to the subset principle also had the least number of tokens!

In this kind of tabulation there is also the methodological issue of going from the idealised linguistic laboratory to the real world, so to speak. These tokens came from real world speech, not experimental elicitation.[35] And, of course,

---

[33]In general, their aspectual classification leaves something to be desired. They allow for the possibility of type-shifting, but as can be seen from the table, since there are only 'plusses', coercion from an activity to an accomplishment is possible by adding a +telic feature, but what about a point-state, such as *It's 10 o'clock* which is derived by the subtraction of a durative feature?

[34]They mention making type counts, but there are no tables for this.

[35]Thus the context that the events (and linguistic descriptions of them) occur in cannot be properly controlled.

categorising verbs by Aktionsart is much easier in the idealised world of the linguist, while determining whether a situation in context is telic or atelic is much trickier. They used a survey method to rate verbs according to aspectual class with minimal frames (i.e. singular subjects and no object, or singular object). Again, it is no surprise that most verbs were rated the same (barring 7.2 percent), but it is unlikely that the situations that the tokens came out of were all of such minimal frames, and there is no indication that the context of each utterance was taken into account to better determine the Aktionsart.

Criticisms aside, the Subset Principle does predict, at least, correctly the *tendency* for the established marking pattern. However, semantic issues involving both event type and (im)perfectivity are involved, it is perhaps better to test the hypothesis using a framework where one can be sure that children are understanding the various tense and aspect markings, rather than possibly using them in a parrot-like way. This challenge is taken up in the next section.

3.5.3   Van der Feest and van Hout

Van der Feest and van Hout are not testing for evidence of the aspectual subset principle as seen above, but rather a variant that combines aspect and tense and put forth an initial hypothesis as to how children acquire tense morphology. The language to be tested is Dutch, which is proposed to have less restrictive tense/aspect marking than the initial child state. Therefore it is predicted that Dutch children will show evidence for the early restricted grammar by restricting their use of tense morphology in certain ways. But, before getting to what the hypothesis is, we shall first look at the tense/aspect facts about Dutch tense morphology (according to (Feest & Hout, 2002)).

The Dutch present perfect functions as a pure past tense, not as a recent past with present relevance as in English,[36] and is compatible with "before today" denoting temporal adverbials:

(4)   a. I heb hem gisteren/net mijn artikel gegeven.
        I have him yesterday/just my article given.
        b. *I have given him my article yesterday.
        c. I have just given him my article.
        d. I gave him my article yesterday.

Notice that when the adverb *gisteren* (yesterday) is used, the only possible translation is with the English simple past. Dutch simple past tense carries no aspectual information, and in particular, contains no coding of completion, unlike the English simple past, which carries completion information at least for some event-types:

---

[36]Ignoring the existential perfect in English.

(5) a. Toen Jan binnenkwam, dronk Marie een glas wijn.
     b. When John entered, Mary drank a glass of wine.
     c. When John entered, Mary was drinking a glass of wine.

Both (5-b) and (c) are possible translations of (5-a), showing that there is no necessary completion information attached to Dutch simple past tense morphology, whereas with the English (5-b), Mary must have finished the glass of wine. In order to allow for the possibility that Mary never finished her wine, the past progressive as in (5-c) must be used. Therefore, Dutch past tense is less restrictive in marking aspect than English.

To summarise, according to van der Feest and van Hout the basic three Dutch tenses are as follows: Present Perfect is past and completion, simple past is just that, and is compatible with both completion and ongoing, while simple present is both now and ongoing. The relevant contrast (they claim) with English is that English simple past is past and completion both.[37] This contrast leads to the tense subset principle hypothesis: "a past tense that carries both Past and completion is more restricted than one that carries simply Past, so the former must be hypothesised first according to the Subset Principle." (p.738)

Importantly, their prediction for the present tense initially indicates only Now and Ongoing. Note that this discounts both the narrative, "hot news" and historical present in both languages, where present tense can be used for remote past as well as having completion implications:

(6) a. In 1066 William the conqueror conquers England.
     b. Henry gets the ball outside the box, he shoots...Oh what a goal!

As will be seen, not taking into account this rather common usage of present tense has drastic consequences for the experiment.

The goal of the experiment is to test three competing theories of tense morphology acquisition: The first being their version of the subset principle, the second, the temporality hypothesis that says simply that past tense form = Past and the present tense form = Now, and finally the Grammatical Aspect first hypothesis which posits that children (learning English at least) take past tense to mean completion and present tense to mean ongoing.[38]

Van der Feest and van Hout use an interactive paradigm pioneered initially in (Wagner, 2001), that will combine completed, incomplete, and ongoing situations, linguistically requiring the children to match these situations to simple present, simple past and present perfect sentences. There is a road that a puppet

---

[37]This isn't strictly true. Past tense states are certainly temporally located in the past, but necessarily 'finished' by speech time. (*I was ill yesterday, and still am.*). When activities and events are used with a habitual meaning, the habit can easily be going on. *When I was a lad I walked 10 km. a day, and I still do. How else would I keep my boyish figure?*

[38]This hypothesis is examined in (Wagner, 2001) which we will see in more detail later in the chapter.

plays along, doing the same (telic) activity at three different locations. The puppet is stopped in the middle of the third location, and another puppet (Cookie Monster) is asked a question using one of the three tense types and must give an answer of a location. The child is then asked if Cookie Monster gets it correct or not.

To illustrate, suppose the puppet finishes a puzzle at the house (the second location) and moves on to the car where she starts to do the puzzle but is stopped in the middle of the activity. Cookie Monster is then asked, *Cookie Monster, waar maakt het meisje een puzzel?*[39] Cookie Monster could answer: "At the house," which is incorrect as the house is now part of a past situation. The child is then asked if Cookie Monster answered correctly. In this case the child must say "no" to get the trial correct. As there are three locations, the first two are past, but the puzzle is completed in one and not finished in the other. The third is ongoing and incomplete.

While this experimental setup does combine tense and aspect information as desired, it does have one rather strange consequence – there is no way to differentiate the subset hypothesis from Wagner's Aspect First hypothesis (which seemed to be the main point in the first place). While we saw that unlike English, Dutch past can be used for ongoing situations, the structure of the experiment has it that even if the activity is not completed at one the first two locations it is nevertheless *not* ongoing when the question is asked. So under these conditions, both hypotheses come down to the same predictions: Children will be able to differentiate past and present morphology only for completed and ongoing situations. For the incompleted situations (where the child gives up and moves on to the next location), the child will not be able to differentiate between present and past tense. They do note on p.744 that the experiment cannot distinguish between the two, and so rely on the justification that the subset principle is more principled.[40]

In order to have a hypothesis that will have different predictions, they have also taken into account the Temporality hypothesis, which simply says that past tense forms mean Past and present tense forms mean Now. Thus, whether or not a situation in the past is incompleted should make no difference whatsoever. Before moving on to an analysis of the results, one more comment does need to be made. As the puppet is carrying out her activities, the experimenter describes what is happening. To illustrate the point I will give the first two sentences of one of the protocols both in Dutch and English translation:

---

[39]This is present tense – *Where is the child doing a puzzle?* Strangely, there is no good simple VP in English to make the activity telic. Doing a puzzle seems to allow pure (i.e. no goal) activity readings.

[40]At the time, the authors would not have known, but experiments conducted more recently by Wagner cast doubt on the Aspect-First hypothesis. However, the results are also of no comfort to the subset principle. This will be clearer in later sections.

> Nu gaat het meisje een puzzel maken. Eerst gaat ze naar de boom.
> Now the girl is going to do a puzzle. First she goes to the tree.

The second sentence is the most revealing as it is in Narrative Present – that is (in this case) a description of something that just happened and completed. But if the children who are being experimented on have their initial values set to Present = Now and Ongoing, they should not be able to understand the protocol, and therefore the experiment is rather flawed. But if they do understand the protocol then they can also understand a use of the present tense that is neither Now or Ongoing. If this is the case then all three of the hypotheses tested are wrong.

Unsurprisingly then, the data elicited from the experiment prove more confusing than illuminating. First off, the adult controls also hesitate to use the simple past for the incompleted situation, but there are periphrastic ways of expressing incompleted or ongoing situations in the past, and this may bias Dutch speakers to using simple past mainly for completed situations (which they do here 100% of the time both in simple past and perfect). Furthermore, they are willing to use present tense half of the time for the completed, past situation. This is against the predictions for van der Feest and van Hout, but we have seen that they do not take narrative present tense into account. Certainly in English if I were asked, "Where does the girl do a puzzle?", I would happily say, "By the tree" even if it was the first of the three locations.[41]

The children exhibit the same behaviour in regard to the present tense, but things get decidedly stranger when it comes to their uses of the two past tenses. Simple past was accepted for 71% of completed situations against 40% incompleted, while the present perfect was accepted for 60% of the completed situations against 42% for incompleted situations. Recall that the prediction for both the Aspect First and Subset hypotheses is that children cannot differentiate in regards to the incompleted situation. However, they only use the present tense for the incompleted situation 15% of the time, much less than either past tense (and much more than the adults would use either past for incompleted situations). The best that can be said is that some of the children follow the path predicted by the subset principle and some do not.

This would seem to be evidence for the Temporality hypothesis in that children seem more likely to use both past tense for a past situation regardless of aspectual information. However, as we saw with the data above, the children are still differentiating between complete and incomplete events in that they are still happier with assenting to a past tense sentence for a completed event. This leads to a picture that has the children using both completion information (the contrast in past tense usage depending on completion status) and temporal information (unwillingness to use present tense for a past incompleted situation) to get at an

---

[41]Again, with "do a puzzle" I would likely also apply it to the incomplete situation as well – I get no telicity here.

understanding of the tense marking. But this is not in the categorical way that any of the hypotheses would have us believe. But, we shall see in the Wagner experiments the same sorts of interactions, and that Wagner herself could only conclude that children are *biased* towards acquiring tense morphology in terms of completion information, not that it is the only factor.

The authors take the fact that there is some differentiation (70% or 60% to 40%) in the usage of past tense relative to a completed or incompleted event to be evidence for the Subset Hypothesis, which is rather strange. First off, with the present perfect the results are 60% for completed and 42% for incompleted. With 24 children and several trials, is this really enough of a statistical difference to make such a claim? Furthermore, if the hypothesis is that the child starts out with a more restricted grammar (for Dutch at least) and then broadens their mapping after hearing counter-examples, then the exact opposite is born out of the data. Ironically, the adult controls seem to have a much more restricted grammar and were loathe to use the simple past for the incompleted situation, while the children were much more likely to. One then wonders where the child is hearing the purported counterexamples.[42]

There is one more quirk in the data that the authors don't bother to mention. The children said "yes" half of the time to the present perfect sentence matched to the ongoing situation, and the same with the simple past in 31% of the trials. This is rather strange as the semantics of the ongoing situation is Now and Ongoing, while the children are predicted to have the same use of the present perfect as the adults do – completed and past. Again, this violates all three hypotheses, but no explanation as to what is going on is given.

Finally, the data would be easier to understand if it wasn't present in aggregate form, but rather divided by child. Perhaps some of the oddities would be clearer if they occurred in only a subset of the children. However, I think it is rather the complexity of the situation being studied that leads one into so many difficulties. To acquire adult competence of tense morphology is not a simple matter as it requires (depending on the language, of course) the combination of temporal, aspectual and lexical concepts to be mapped onto the appropriate syntactic forms. "Deviant" uses of tenses such as the Narrative or Historical present must also be learnt. It is therefore no surprise that the results both here seem to show a variety of factors interacting. Indeed, a more correct approach would be to examine the process involved, and allow for a number of cognitive factors to be harnessed in this process. In this spirit, the next work from Weist et al, which brings universal cognitive development into the equation.

---

[42]The authors do acknowledge the oddity in the data and propose using a 'when' clause along with explicitly saying that it wasn't finished, in order to tease the aspectual neutrality into the Dutch simple past.

3.5.4 Weist (1991)

In the last section we saw an attempt to justify experimentally a generative approach to language acquisition, specifically tense morphology. In this section, we shall also look at experimental evidence to justify a certain hypothesis regarding the acquisition of the same morphology, but from a different point of view. Rather than a single language specific learning mechanism such as the Subset Hypothesis, (Weist, 1991) proposes a general cognitive development process, in which children will universally develop the same temporal conceptual system in a number of stages. With this system as a basis, the child is then able to learn the target languages temporal system in a more piecemeal manner. Along with the general and universal cognitive system, the fact that different languages have vastly different temporal systems allows for many language specific variants to also come into play. Thus, the predictions as to what children learn when will also depend on the language they are acquiring. In a sense, this is a kind of bridge between the more 'universal' approaches (be they generative or cognitively based) and the language specific ones such as those of Bowerman and Tomasello.[43]

Background

Weist's account of language learning fits into a neo-piagetian paradigm, where the gradual development (language independently) of first spatial and then temporal systems, allows the child enough foundation to then learn their language's respective space and temporal grammatical systems. We shall ignore the parallel account of the development of spatial concepts here, and concentrate solely on Weist's account of the child's development of temporal concepts. We shall look briefly at the 4 stages and for ease of the reader's comprehension show their analogue in language. After this will come an account of a cross-linguistic study to test the final three stages.

Weist proposes a developmental progression of four different temporal systems, the last of which corresponds to the adult temporal system. The process itself involves the sequential emergence and then integration of the Reichenbachian concepts of Speech Time (ST), Event Time (ET) and finally Reference Time (RT).

The first system that the child develops is called the Speech Time system, due to the fact that child language at very early stages is limited to speech-time coding only. Linguistically they do not code tense, aspect or modality. Nevertheless they can make reference both to absent objects as well as prior experiences and anticipated future situations.

From studies of a number of languages, between 1;6 and 2;0, children begin to

---

[43]Note that language specific does not mean Skinnerian Behaviourism, only that the universal mechanisms (e.g. intention and goal recognition, pattern matching, general learning) are more general than even those proposed by Weist.

express a deictic relationship between ET and ST, for example being capable of making a past/non-past distinction, and in languages such as the Slavic languages, make an imperfective/perfective aspectual distinction. This signals the entry into the second stage, the Event Time system, where the conceptual analog is that children have a grasp on the distinction between ST and ET, but keep their temporal point of reference always at ST (i.e. no RT yet). Conceptually this requires the child to be able to retrieve an episodic representation, and determine its relationship to the moment of speech.

Around the age of 2;6 to 3;0 children begin to use temporal adverbials or adverbial clauses to establish the temporal viewpoint, RT. In languages that have it, children also will relate ET to RT with a relative-absolute tense such as the English present perfect (where ET is before ST, but RT=ST). Notice that while all three components of tense have now come into play (ST, ET and RT), children are still limited to relationships between two temporal intervals only. For example, while the English present perfect does have all three components the 'present relevance' meaning of it forces RT to be identical with speech time, and the temporal distinction is between an event in the past and its relevance for now. For this reason, the third stage is called the Restricted Reference Time system (or RRT).

Finally (from around 4;0 to 4;5),[44] children begin to develop more flexibility in constructing temporal configurations, and begin to use adverbs such as 'before' and 'after', and in languages such as Finnish begin to use the past perfect. Since use of the past perfect presents ET, RT and ST as independent, the child must be able to co-ordinate three separate temporal intervals (*I had already made dinner when Mary entered the house*). For this reason, the final temporal system is called the Free Reference Time system (or FRT).

Conceptually, we can see a clear developmental path that the child must take. After the first stage, the child must be able to handle some event representation and determine its relationship to the absolute time point of Now. Acquiring RT requires the child to have the capacity to establish a referential context (or temporal perspective). Initially, this viewpoint is tied to speech time, but is then allowed to free itself, requiring the child to now put the location of some episodic representation relative to this context and to Speech Time.

Armed with this conceptual foundation, on the linguistic side the child must figure out how temporal information is coded in the target language, and then construct a system of rules to form the basis of comprehension and production of various temporal configurations. Moreover, the child must also develop (understand) a lexicon of temporal words to deal with conventional time. In this paper, this is really all Weist has to say regarding the linking of the conceptual development to language acquisition, which does not seem to be enough to accomplish the

---

[44]This 'finally' seems to be the most language specific of all stages, with evidence that Bulgarian and Finnish children move to this stage rather early.

linking process. In (Weist, Lyytinen, Wysocka, & Atanassova, 1997) he appeals to Slobin's (self-refuted) Operating Principles (Slobin, 1985),[45] but one wonders if the more general, developmental learning accounts (Slobin, 2001) can't handle the linking problem if the conceptual apparatus is in place.

With this background in mind, we shall look at tests conducted by Weist and associates on American, Polish and Finnish children aged from 2;6 to 6;6 designed to test the existence and timing of the latter three stages, as well as account for language specific interference in the move from conceptual stage to a learning a languages specific temporal system.

The Experiments

(Weist, Wysocka, & Lyytinen, 1991) conducted six total comprehension experiments, with two experiments meant to correspond to each of the last three developmental stages. All experiments involved a forced sentence to picture matching task, where children were shown two pictures and are then given a minimal pair of sentences and asked to choose which sentence matched which picture (only the first choice being scored, of course).

The tests corresponding to the Event Time system involved a test of Absolute Tense (Past vs. Future) as well as comprehension of the distinction between perfective and imperfective aspect. For example, in the Tense Tests, there would be a pair of pictures such as two children in the snow. In the first one, a child is being hit by a snowball, with the other child is at the end of their throwing motion (now *sans* snowball). In the second, a child is holding a snowball ready to throw it at his concerned companion. The subjects would then be presented with the pair of sentences *The girl (threw/will throw) the snowball*, where the first sentence matches the first picture described, etc. Both the Polish and Finnish minimal pairs involve a verbal Past/Non-past distinction, where in Polish the verb is perfective in both sentences, and in Finnish the object (*snowball*) is in the accusative case in both sentences).

The aspect tests involved a pair of pictures such as a girl busily drawing a flower and the same girl smiling proudly over the picture of a finished flower drawing. The English minimal pair was *The girl (drew/was drawing) a flower.*[46] The Polish variant involved perfective and imperfective marking on the verb for draw (*rysowa*), while the Finnish aspectual contrast was achieved by case-marking on the object – *drew* corresponds to accusative case on the object, while *drawing* corresponds to partitive case on the object.[47]

For these two experiments, they predicted that all children should be able to

---

[45]A kind of conceptual nativism.

[46]Note that while the first could also be described as "The girl *is* drawing a flower", tense needs to be kept constant and aspect allowed to vary in order to get a true minimal pair.

[47]This will be looked at in more detail in the section on Wagner's final experiment.

handle the Absolute Tense test by age 2;6, and that both Polish and American children should correctly do the Aspect Test. But since Finnish codes aspect in a number of ways (verbal morphology, 3rd infinitive and case-marking as used here) and case-marking is used for many other functions that the Finnish child is also learning, the authors expect a slight delay in competence for the Finnish children. Furthermore, given the fact that aspect is such a fundamental component of Polish verbal morphology, they predicted that Polish children would have a slight advantage over the Americans, but this acquisition is, anyway, likely to have occurred by the age of 2;6.

The second set of tests for the RRT system involve the contrast first between a remote and immediate adverb, as well as a contrast between the present perfect (with *just*) and the simple past plus adverb. For example, the subject is presented with two pictures of a girl thinking (her thoughts shown in a cartoon thought bubble). In the first she is at bed praying, and the thought bubble shows the same girl dancing. The second picture shows the same girl eating an ice cream, but also with a thought bubble of the girl dancing. The minimal pair is *The girl will dance (tomorrow/in a while)*. Thus RT is tied to ST and the child must be able to somehow coordinate this with the context provided by the thought bubble. They also must have a lexicon for time good enough to distinguish between the two adverbs. The minimal pairs in Polish and Finnish are no different linguistically.

The second contrast involves two pictures such as an excited family examining a Christmas tree in their living room, and a picture of the same family eating dinner with the Christmas tree off in the background. The English minimal pair is *Father (has just brought the Christmas tree/brought the Christmas tree yesterday)*. Thus if the subjects have mastered the RRT system, they should be able to differentiate the present perfect (or its equivalent) from the simple past. Finnish has a present perfect similar to English, while Polish must make do with the same perfective past verbal morphology plus the differentiating adverbs (*just* and *yesterday*).[48] The authors predict a slight advantage for the Finnish speakers over the American speakers in this test as Finnish children do begin using present and past perfect before American children. However, the addition of adverbials along with the present perfect/simple past contrast may negate this. Furthermore, this fact about Finnish children may be a sampling accident anyway, as the comparison is with American children. Studies of Scottish children show that they acquire the perfect more rapidly than their English or American counterparts. (Gathercole, 1986) concludes that a contributing factor is that the perfect is much more widely used in Scottish English,[49] and hence the children are much more exposed to it.

The final tests (for the FRT system) consisted first of a contrast between

---

[48]Which given the test in the paragraph above, seems to make this test rather redundant for the Polish speakers.

[49]Including in speech to children.

'before' and 'after' and also of a contrast between the simple past and the past perfect (both of which require yet more sophisticated temporal co-ordination). The pictures for 'before' and 'after' would have things like a woman in the bathroom on the telephone. In the first picture, a bathtub is filled and the floor is clean and dry (and the woman is fully clothed). In the second picture, the woman is wrapped in a towel, the bath empty, and the floor filled with puddles of water. The minimal pair involved is *Mother answered the phone (before/after) taking a bath.* There are no interesting linguistic distinctions for the Polish or Finnish versions.

The simple past vs. past perfect distinction would have a pair of pictures such as a boy lighting a fire while another boy is walking into the picture. The second picture has the fire already lit when the second boy arrives. The English minimal pair is then, *The boy (started/had started) a fire when his friend arrived.* Finnish, having its own past perfect, works analogously, but the Polish minimal pairs instead has to substitute 'before' with the past perfective to get at the intended meaning. Again for Polish, this boils down to something like a before/after problem, but given that the child has to co-ordinate two events relative to each other and speech time, they would still need a flexible temporal systems to handle the complex configuration. Again, the authors expect the Finnish children to do solve the past perfect/simple past problem a little earlier than the others.

Results:

The results, as one might expect, were varied. As predicted both the Polish and American children passed the ET system problems by age 2;6, but neither group showed full mastery of the RRT and FRT systems until 5;6 or 6;6. In fact neither the American or Polish children managed to show mastery of the past-perfect/simple past distinction, even at age 6;6.[50] Nevertheless, some of the children from these groups managed to pass the more complicated tests by 3;6 or 4;6 (in fact, some 2;6 Polish children managed to do quite well on the before/after test). Weist et al interpret these results as showing a transition period at around 4;6 (which is consistent with their elicitation tests), which signifies that the process of what they call de-centration is well under way. However, mastery of a languages temporal system does take a bit of time, which accounts for the 2-year lag. Moreover, while there is a significant distinction between an ET an RT system, they were unable to find any way to distinguish between RRT and FRT systems. Finally, this transition between ET and RT parallels a slightly earlier transition in child's spatial conceptualisation.

On the other hand, the Finnish data yielded some surprises. While there was a prediction that the Finnish children would lag behind a bit on the aspect test, there was no prediction that it wouldn't be until 6;6 that children showed mastery

---

[50]This is never explained.

of any of the tests, including both ET system problems. Ironically, they were the only group of children to be successful with the past perfect problem. As briefly hinted at before, the use of object marking for aspect was expected to hinder the children's mastery of the ET system, but not so drastically.

Tense was not expected to be effected and the authors proffer a variety of explanations including the fact that unlike English and Polish there is no clear tripartite (i.e. Past, Present, Future) distinction in Finnish, but only past/nonpast. Thus the sentence using non-past morphology could be ambiguous between meaning now and some future time.[51] But, there are some problems with this explanation. First of all, English *will* is not simple future marking, but modal and can include present time reference as well (*That'll be John ringing the bell*). Furthermore, even if the sentence (with non-past morphology) is interpreted by children as having present tense meaning, there is still only one correct match to the two pictures (at least in the snowball case). There is one picture where the child is already hit by the snowball – this matches neither present or future meaning. But the picture of the child, holding a snowball in his hand and arm in a throwing motion can be perfectly described in English present tense (at least) as *the child is throwing a snowball*.[52]

With the remarkable failure of the Finnish children at the ET system, Weist and colleagues ran a second test with the same pictures, but this time with additional minimal pairs. For the past/future distinction, the non-past sentence was changed to *intends to*, as in *the boy intends to throw a snowball*. For the aspect test, a 3rd type of Finnish infinitive is used to get the aspectual distinction instead of partitive vs. accusative case marking.

Both revised tests had the Finnish children exhibiting some improvements – mastery of both now dropped to 4;6. This still isn't the 2;6 seen in both American and Polish children, however. The authors attribute this to the complexity of the verb forms. The addition of the 3rd infinitive form may indeed add to the complexity. But, this does seem rather dubious for the condition of the experiment meant to get at tense, where the only change is a non-past form of an ordinary verb.

But notice the common factor in both of the changes – children have direct information regarding the subjects' intentions. With the future tense sentence, *intends to X* can only have a future reference for X – it is implied by what an intention or plan is. Switching the aspectual coding from the object to the verb

---

[51]Other explanations include complications in morphology in spoken Finnish, but these are rejected as English has the same issues (irregulars and slightly different phonological endings depending on the verb), and the possibility that Finnish children are developing a more complex system more slowly.

[52]Recall the event schemata presented earlier in this chapter. Conceivably, the arm motion with snowball in hand can be seen as one stage of the event, and thus present tense is appropriate. On the other hand, one could also identify *throw* with just the last stage, bringing about future time reference.

may also allow children access to intentions by the connection of the verb to the subject/agent.

A closer look at the experiments in general shows that the children need to be able to recognise intentions, goals, and make inferences to succeed in the test at all. For example, the adverbial test (*yesterday* vs. *in a while*) has a girl with a thought bubble imaging herself dancing. This has to be understood for the minimal pair to make sense at all – the inference made of a girl about to go to bed and a girl eating ice cream helps the child determine the appropriate adverbial. Similarly, the past perfect test shows excitement of the family at the Christmas tree versus the family eating dinner and ignoring it. Knowing that the family is excited because of the Christmas tree allows the subject to infer present relevance, and hence, appropriateness of the perfect.

In fact, every single pair of pictures shown to the subjects involve a human engaged in, desiring to engage in, or having engaged in some activity or other. This might lead one to think (along with the anomalous Finnish results) that Weist et al may be overlooking an important cognitive factor – that even the youngest of children are able to recognise intentions and goals of others with astonishing accuracy. This point did indeed occur to Wagner, who looked at what happened to English children's acquisition of the event-time system (more specifically the aspect part) when the subjects are deprived of agents or intentional information in the experimental set-up. After examining Wagner's basic modification of Weist and testing for the Grammatical Aspect First hypothesis, we will then see how understanding of others' intentions can effect comprehension of tense and aspect.

3.5.5 The Wagner Experiments

(Wagner, 2001) is an account of two experiments designed to test what she calls the *Aspect First Hypothesis*. It is yet another theory to account for children's acquisition of grammatical tense. Stated simply, it is the hypothesis that children initially use verbal, tense morphology not to mark semantic tense, but rather grammatical aspect. Of course, we have seen that in languages such as Spanish and French that verbal morphology actually may mark both tense and aspect with a single morpheme, or in the case of the English *-ing*, aspect alone. Given this conflation, Wagner attempts to design the experiment to show that children (up to age 4 or so) are sensitive only to the aspectual meanings of the morphology, and though even children under two can productively use tense morphology, they are not encoding tense information, but rather are taking it as encoding completion information about the event.

One difficulty is that in English,[53] there are often completion implications for uses of the simple past. Therefore, in order to test the hypothesis that young children are encoding *only* completion with past tense morphology, the linguistic

---

[53]We already saw this is (Feest & Hout, 2002).

tests must be carefully set up to rid the past tense of completion entailments. Instead, if children will consistently associated completed events with *was V-ing*, and incompleted[54] events with *is V-ing* (regardless of temporal order), there would indeed be evidence that children are indeed associating the irregular copula morphology with completion information and not time. Indeed, Wagner tests children with examples of past and present progressive sentences, both of which carry no implications regarding completion.

3.5.6   Wagner Experiment 1

The first experiment is a (fairly substantial) modification of experiments done by (Weist, 1991), which were claimed to provide support for the *temporality hypothesis*. Children were shown two pictures, given two sentences and asked to choose which picture matches which sentence.[55] For example, a child would be presented with a picture of a boy about to throw a snowball and a picture of someone being hit by a snowball. The child would then have to match the pictures to the sentences *The boy will throw the snowball* and *The boy threw the snowball*. Children as young as 2;6 performed better than chance on the task, thus apparently showing that young children *do* associate tense morphology with temporality.

Wagner correctly points out that a major problem with the above study is that it contrasts past and future, leaving the present tense untested. Analogically to the above discussion regarding past tense and completion, the English 'future tense' is much more than a temporal marker but also a modal one, in this case a marker of irrealis. So, if children can understand the realis/irrealis distinction they would perform better than chance on the task, even if they did not associate temporal ordering with the tenses as of yet.

Therefore, Wagner includes all three tenses for the task, providing the proper contrast in order to find out the extent of completion influences on tense interpretation (children need more than a realis/irrealis distinction to succeed). Moreover the experiment includes both telic and atelic events in the example sentences to serve as a test for the *lexical aspect first* hypothesis[56] and localist hypothesis.[57] If comprehension follows production (where children use *-ing* far more with atelic events), then there should be some advantage to sentence to sentences with atelic predicates. Finally, the test is done with an interactive paradigm, depicting acted out events rather than static pictures. The purpose of this is to provide children with an experimental environment closer to the way they actually expe-

---

[54]My apologies for this awkward word, but I am trying to contrast not-yet-complete (i.e. ongoing) with events that stop before reaching the natural endpoint.

[55]Of course, only the first choice can be scored, as the second choice is then trivial.

[56]This is the view that perfective and imperfective marking is used to mark telicity.

[57]This is a verb by verb learning hypothesis, such as the verb-island hypothesis of (Tomasello, 2003).

rience events in time. With this in mind, it is probably easiest to understand the changes made by describing the experiment.

The setup involves a picture of a road drawn on a piece of paper. The children are then introduced to a toy kitty who has a penchant for performing various activities on the road. With each trial, the kitty goes along the road performing the same event three times (at the beginning, middle and end of the road). An inky rubber stamp at the bottom of the kitty allowed the children to trace kitty's path, providing a cue to proper temporal ordering of the events. The trial consisted of first the child watching kitty performing the event at each location (to completion), and then a second run. In this run, in the middle of the second event the child is asked either *Show me where the kitty was V'ing*, *Show me where the kitty is V'ing* or *Show me where the kitty is gonna V* – corresponding to past, present and future respectively. Note that although the second location (being still in progress) is compatible with all three sentences, a correct response to the present tense query was considered to be the second location only. Furthermore, there were control trials where the query was supplemented with the appropriate temporal adverbial (e.g. 'before', 'now', 'in a second'). Thus the past tense query would then be *Show me where the kitty was V'ing before*. This control condition serves as a reminder of exactly what is being tested. The aspect first hypothesis says nothing about children's development of temporal notions *per se*, only about their acquisition of temporal morphology. It is entirely compatible linguistically for children to initially put all their temporal ordering information in the aforementioned adverbials, reserving temporal morphology for completion information.

There were a total of 12 distinct conditions for the test – the three tenses, telic or atelic, and with or without temporal adverbials. There were three different telic and atelic events, including drawing a face and emptying out a cup (telic), and playing with a toy and hopping about (atelic). For a concrete example of a query, with the drawing a face event a sample of the queries would be *Show me where the kitty is drawing a face* and the control, *Show me where the kitty is drawing a face right now*. It should also be noted that children of age groups of two and three years of age were compared.

The results do show that even two year olds perform better than chance (though somewhat marginally) when asked to point out a location as an answer to a given query. The 'past' location was chosen most often for the past tense queries, etc. Three year olds did rather better than the two year olds and both groups performed significantly better on the control adverbial queries; however, the difference between the test and control queries for the three year olds was not nearly as prominent as it was with the two year old children. Also, there was no difference in regard to lexical aspect, i.e. telicity did not seem to have any interaction effects.

On first glance, the results seem to confirm Weist's original findings – children do understand tense markings, at least as when carried by the copular auxiliary,

as conveying semantic tense. But, though linguistically care was taken to separate tense and completion implications (by use of *is/was V'ing*), the events as presented did not mix things up enough. All of the telic events in the past situation were, in fact, completed, while the atelic events came to a natural end (i.e. no sign of an interruption). So, it could actually be the case that children were understanding the past progressive as indicating a completed event and the present progressive as indicating an ongoing or incompleted event. Therefore, a second experiment was done to test this possibility.

3.5.7   Wagner Experiment 2

The setup is the same as in the first experiment, with a road drawn on a piece of paper and an inky kitty performing various activities on it. In this case, the kitty performs only telic events, and there are only two telic events per trial, representing past and present. But the major difference is that in half of the trials, kitty fails to complete the telic event at the first location, giving a novel situation of past and incompleted. This is important in differentiating Weist's temporality hypothesis and the aspect first hypothesis. The temporality hypothesis predicts that since children have an adult-like understanding of tense morphology, whether or not the past event is complete is irrelevant to the child's comprehension of the queries. In fact, given the pragmatics of the use of a past progressive rather than the simple past, one might argue that performance would improve with a past, incompleted event. In contrast, the aspect first hypothesis predicts that since the child is using completion information to understand tense morphology, they will run into problems if both the past and present event were incompleted or ongoing. Another significant difference from the first experiment is that the adverbial controls were omitted.

Here, the results (which now include a group of four year olds as well) are rather interesting. The group of two year olds still managed to successfully distinguish between the past and present queries in the half of the trials that involved completed past events. But when presented with the trials with the incompleted past events, the two-years were completely lost. Not only did they fail as predicted when presented with the past-tense query, they also did significantly worse with the present tense query. The three and four year olds did better (both above chance, with four year olds doing much better), but even the four year olds saw a performance drop when presented with the incompleted past event.

3.5.8   Conclusion

Combining the results of the two experiments, Wagner concludes that there is limited support for the aspect first hypothesis. On the one hand, the two year olds were completely lost when the completion information was switched about, which does suggest that two year olds (and to a lesser extent, slightly older children) do

use completion information to interpret tense morphology. And if we look only at the results of the second experiment, it seems that the two-year olds are not using temporal ordering information at all.[58] But, Wagner does acknowledge that the results of the first experiment are much too strong to take completion calculations as the *only* factor in the initial stages of tense morphology acquisition.

She brings up two possibilities as to reconcile the apparently conflicting results of the two experiments. First, it could be the case that children are conflating tense and grammatical aspect on the same morphemes, and when (like in Experiment 2) they are presented with a tension between the time and aspect, havoc results. This is not altogether unreasonable, as we have seen that many languages do conflate tense and aspect on the same morpheme (e.g. many romance languages), and even the language that was tested (English) has some conflation of simple past and completive aspect. Secondly, children may be using the completion cues to help with temporal sequencing in a similar way that causal information is used. (e.g. the classic *Max fell. John pushed him.*).

3.5.9   Wagner Mark II – Intentions into the Picture

Before examining another factor that can be a participant in the tense/aspect acquisition process, a brief summing up is necessary. (Weist et al., 1991) conducted a series of experiments to find correlations between his neo-piagetian developmental model and acquisition of spatial and temporal linguistic systems. The stage that concerns us here is the second stage – the so-called event time stage. This applies to children around 2;6 and holds that at this point in development, children can distinguish between Speech Time and Event Time as well as comprehend the perfective/imperfective distinction. This requires that children can both conceptually distinguish present time from past time, as well as events in both the internal and external perspectives. Moreover, they have also mastered enough of their language's tense/aspect distinctions to express and understand these distinctions linguistically.

Their tests for the perfective/imperfective distinction involved a forced choice sentence-to-picture matching task. Children were shown two pictures – one a picture of a completed event (e.g. a girl smiling next to a finished drawing of a flower) and the other a picture of a yet-to-be completed version (here a girl busily drawing a flower). The children were then given two sentences describing the pictures, one each of a perfective/imperfective minimal pair (with tense held constant). Here the sentences would be *The girl was drawing a flower* and *The girl drew a flower*. Children were then asked to match the sentences to the picture. It was found that children as young as 2;6 would reliably match the imperfective sentence to the yet-to-be complete picture and the perfective sentence to the finished picture, just as is predicted by the event-time developmental stage.[59]

---

[58]It's a pity that Wagner omitted the adverbial controls for the second experiment.

[59]Of course, the imperfective sentence is compatible with both pictures, but hopefully the

The experiment was translated and carried out as well in two languages very different from English – Polish and Finnish. The Polish children performed the same as the English children, successfully matching the sentences to the pictures. On the other hand, the Finnish children performed abysmally, not passing the test until age 6.[60] A reason for this (noted on by Weist) might be due to the way Finnish marks aspect. Instead of verbal morphology (as in English and Polish), Finnish marks imperfective and perfective aspect (again, really completion information to an extent) on the object:

(7)   Maija luki kirjaa
        M.   read book-part.
        Maija was reading a book.

(8)   Maija luki kirjan
        M.   read book-acc.
        Maija read (all) the book.

Thus an object marked with partitive case marking gets interpreted as imperfective and an object with accusative marking gets interpreted as perfective. This suggests that it is the accident that aspectual marking comes on the object rather than the verb in Finnish that makes it harder for Finnish children to acquire. For example, a study by (Hout, 1998) suggests that children (English and Dutch) have difficulties in the area of the count/mass noun distinction. Like the Finnish case marking, the count/mass noun distinction can change the completion entailments of a given predicate – compare *John drank beer* with *John drank a glass of beer*, only the latter of which is telic. Therefore, if it is difficult to connect the object to lexical aspect, it is not so surprising that Finnish children find it difficult to connect grammatical aspect to object marking. In fact, when Weist re-ran the picture experiment with sentences containing another way to code aspect (via a special verb form) the Finnish children did much better (though still not as good as the English or Polish children).

Wagner sees the problem as perhaps more conceptual. Instead, it may be difficult for young children to get aspect information purely from the objects themselves. That is, how well can children relate an event taken as a whole with its completion entailment in the world? Finnish happens to mark the completion on the object, making the child's learning task comparatively difficult. But English and Polish children may be able to use the other facet of telicity to help them to get the proper mapping – that of intention or goal. Recall that when the English progressive when applied to accomplishments was discussed in Chapter 2, there

---

forced-choice nature of the task makes the distinction clear.

[60]In fact, they performed worse to some degree or other on 7 of the 8 tests in comparison to the Polish and English children.

were two important elements – a partial object on its way to completion, and a goal or plan to accomplish this. Having verbal morphology mark aspect (as in Polish and English) provides a connection to the agent/subject which the children then might use to infer a goal or intention to have the object completed. Earlier in the chapter (when discussing the psychology of events) we already saw that children as young as 1 year old are quite good at inferring intentions. Indeed, understanding of intentions may be one of the ingredients used in the process of tense/aspect acquisition.

What is then needed is to see is how the English children fare when given only object information, and no way to discover intentions or goals. The pictures used by Weist were intrinsically agent oriented – one showed a girl busy at work drawing, the other showed her happy and holding a finished drawing. Perhaps the little girl was enough for the children to infer the different goals of the girl and link the morphology to her goals. Instead, (Wagner, 2002) designed an experiment with the aim of testing English children on the same sorts of sentences, but this time with visuals that lacked any agent-oriented information.

Instead of pictures the children are presented with pairs of toys depicting different stages or versions of the same event(type). For example they would be presented with two puzzles, one completed filled in and one half-filled in, or an empty-cup and a half-empty cup. The twist here is that, unlike the previous interactive experiments the puppets perform the activities behind a screen and the children can only hear them talk about what they are doing (in order to connect each puppet to the event they did). When finished, the screen is lifted revealing two versions of the event, say a finished puzzle and a half finished puzzle. The characters, now separated from the events, would utter an imperfective or perfective sentence (depending on which event they did), such as *I filled in a puzzle* or *I was filling in a puzzle*. After the sentences were repeated a few times, the child then had to match each puppet with one of the objects. Naturally, only the first match was scored as the second is forced. A match was considered correct if the imperfective sentence was matched to the incomplete object or the perfective sentence was match to the completed object. There were also adverbial control trials where the test sentences directly specified the aspectual information (e.g. *I'm partly/all done/in the middle of/completely*. Finally, a tableaux test (object picture and sentence matching) was done with adults as a further comparison.

In a reversal of Weist's findings, the 2 and 4 year old children performed no better than chance on both the perfective and imperfective sentences, and it isn't until 5 years old that children even perform above chance with the perfective sentences. Five year olds perform as well on events with perfective sentences (nearly 100% right), but are still at chance with the imperfective sentences. Both results are rather interesting but for different reasons. It does indeed seem that when denied agent information that children under 5 are at sea in regard to how to match the sentences to objects in varying stages of completion. This supports the idea that children are using a strategy of intention/goal reading in order to

figure out the completion entailments.

The five year olds do show that they have mastered the art of linking object information alone to aspect. They did perfectly on matching the perfective sentences, but seemed to have no preference with the imperfective sentence. This seems not to indicate so much that the 5 years do not understand the completion entailments, but don't understand the pragmatics of the task. The imperfective sentence is semantically fine with either the completed or incompleted object. It is only the nature of the task itself that forces the match of the imperfective sentence to the incomplete object. However, the five year olds may not be checking the sentences against each other (and the objects), perhaps not having mastered the adult discourse game.[61]

Wagner's explanation of these results is that children's starting strategy is to start the mapping process with concepts they understand. Here there will focus on intentional properties in the language, since they are well-equipped to deal with this conceptually from quite an early age. Perhaps after hearing (for English) progressive sentences that are devoid of an intentional component such as *The water is boiling* or counter to intentions (*Mommy was falling down*), they slowly expand their meaning of the progressive to include focus on object completion.[62] Indeed, this kind of piecemeal learning of tense/aspect morphology presents a different kind process to the generative view presented earlier. Here the child systematically builds up a grammar that is not yet based on adult principles; she uses the tools she has available along with language specific environmental input to slowly reach adult competence.

3.6   Preliminary Conclusion

Thus far, we have seen a number of experiments dealing both with the psychology of events and event segmentation and then experiments regarding children's acquisition of tense and aspect morphology. It is quite a lot to digest, and, especially in the case of the child experiments, it is reasonable to ask if anything can be concluded from the apparently contradictory results of the various experiments.

First there is the obvious answer that learning tense/aspect morphology is a lengthy process that only *begins* when the child begins to use it a productive way. Furthermore, it also seems that trying to tie this aspect of language learning to one simple principle (i.e. subset principle, aspect before tense, etc.) is more than likely a futile effort. Instead, it is far more likely that children are bringing to bear all of the cognitive resources at their disposal to master these complex systems. Given the complexity of the task, it should hardly be surprising that the young learners will use everything from intention and goal recognition to

---

[61]In *single sentence* tests, adults perform identically with children on imperfective sentences.

[62]There is a hidden assumption that children do not have the conceptual capacity to connect objects to events until about age 5 – this seems reasonable enough.

pattern matching and statistical learning to basic spatio-temporal segmentation to be able to both understand (conceptually) and communicate (linguistically) time and events.

Indeed, it is really this last point that should be taken from the above discussion. Too often, though perfectly understandable as a starting point, the mapping problem between form and meaning that the child has to tackle is far too idealised. For example, in the acquisition experiments it was the child's task to match statements about events to experimental situations. This could be a perspective on the event (aspect) or locating the event in time (tense). The difficulty in these experiments was thought to be the mapping from the linguistic form to concept (temporal or aspectual).

If Chapter 2 detailed the enormous typological diversity and complexity of the concept 'event', this was only reinforced by the discussion in the first half of this chapter, which detailed the psychological complexity of this same concept. But, in the experiments seen thus far, no one asked the question as to whether the children actually had the same concept of 'event' as their adult counterparts. If the answer is no, then it would seem that this fact alone would profoundly effect any experiment, in the sense of what the children are mapping to with their language may not be what is expected by the experimenter. Also, we begin to see a picture of the child developing both its language and concepts simultaneously, further compounding any effort to find a simple categorical principle that would be the holy grail of language acquisition.

In fact, (Wagner & Carey, 2003)[63] conducted just such an experiment as to how children individuate both objects and events.[64] It turns out that three to five year old children do not segment events in precisely the same way as adults, and rely more on spatio-temporal differentiation than intentions or goals. Again this is not a categorical distinction – children as young as 3 have some grasp on the telic/atelic distinction and the linguistic form presented (if goal-based) does cause them to individuate the event on goal-based criteria. Nevertheless, they are far more likely than adults to be biased toward a spatio-temporal method for individuation (while adults are biased the other way round). What appears to be happening is that the child is using the same basic cognitive ingredients that the adult does in individuating events, but only in different proportions. This may seem an innocent distinction, but could have implications for ways in which the children use and understand aspect as well as how well they can deal with coercion effects (that often crucially depend on how the event is segmented).

The notions of tense and aspect used in the experiments also suffer from an idealisation that needs to be moved away from if we want to have a greater understanding of the acquisition process. All of the experiments involved some sort of variation on a minimal pair of *single* sentences that are to be matched to

---

[63]Which shall be examined in greater detail in the next section.

[64]Recall the parallels between the two in the beginning sections of this chapter.

situations created by the experiment. The only criteria for getting an imperfective sentence correctly matched is to match it to an ongoing situation, and the criteria for tense was to match the target sentence to the temporally appropriate situation. However, both of these construals are only part of the story.

We have already seen criticism of some proposed mappings for being far too narrow (e.g. van der Feest and van Hout ignoring that fact that present tense has narrative uses and past time reference possibilities, but then using these in the protocols read out to the children). But there is a larger issue here. While it is now more or less widely accepted one cannot see the whole of syntax or semantics by looking at sentences in isolation, this is especially true when looking at tense and aspect. Only in the context of a wider discourse, does the full power and complexity of these systems come into play.

For example, it is not just that imperfective morphology indicates something ongoing or (potentially) incomplete, it also has discourse functions such as serving as a background or temporal anchor point for following (or perhaps preceding discourse) – *It was raining that spring morning. John dutifully took out his umbrella.* Tense, as is now well known, is not simply the location of a single event on a timeline (relative to speech time), but also is strongly connected to an ongoing discourse. Thus a simple past tense statement out of context, like *Chapman breathed a sigh of relief*[65] is rather strange. Certainly the listener would put this event before utterance time, but really not know what to do with it. Only with preceding discourse, adverbials or given context does it make any sense, e.g. *When the doctor told him that the tests were negative,...*[66] Arguably, a child cannot be said to have mastered her language's tense/aspect system until she has mastered not only the tasks we have seen thus far, but also the more involved discourse uses of the system. Certainly, if we are interested in how a child learns a particular language, these are just as much the 'meaning' of tense and aspect as simple localisation in time or basic perspective taking.

With this in mind, it is a good idea to extend the experiments seen in this chapter to a multi-sentence discourse, to see at what age they can use and understand tense/aspect combinations as ways of temporally ordering a multi-faceted discourse. Animations or puppet performances could be done regarding the standard situations (putting on clothes in cold weather, going upstairs to room and closing door, driving, etc.) What would be needed are events that could be conceived in either ordering (so driving too fast and getting a ticket would be out). This could handle tense and causality ordering (i.e. forward or backward sequential or simultaneous) – the child would then be given a minimal pair of, say, two perfective sentences and have to match them to the two situations (which would be opposite temporal orders, of course). I think this could be difficult since per-

---

[65]Example from (Steedman, n.d.).

[66]Note that the first sentence would work rather well as the first line of a story. Here the reader would be disposed to ignore the lack of context for a bit, in fact adding to the tension of the story.

fective temporal ordering is so dependent on world knowledge and context – great care would be needed to come up with unambiguous situations.

For testing aspect (possibly a bit easier), two events are needed – both of which could be used alternately as a background and a main event. The minimal pair would then be to alternate perfective/imperfective marking and to see if the child could correctly match the mini-discourses to the situations. For the older children we could perhaps alternate between simple and absolute relative tense over a two sentence discourse. Can children handle a discourse with a present tense first sentence and past-perfect follow-up and correctly order the events (and vice-versa)?

Perhaps, instead of (or in addition to) two minimal discourse pairs the situation could be presented and the first sentence given. The child would then be given two sentences to choose from to continue the discourse (again minimally so only either tense or aspect, but not both vary). There could also be a 'production' exercise where children are given the mini-discourse and then either draw, play-act, or play with toys to illustrate the situation.

3.6.1 Boland

There is one more set of experiments by Wagner to examine. But, first there will be some brief comments about a recent PhD thesis (Boland, 2006) that has, unfortunately, has been published too recently to adequately do it justice. Nevertheless, a few remarks will be given and the interested reader, especially as far as it is germane to what has been explored in this chapter.

(Boland, 2006) is an astounding study about tense, modality and aspectual systems (TMA systems). It investigates the connection between limits on the variation of these systems cross-linguistically and stages of language acquisition. Cross-linguistically, she proposes a typological hierarchy of three classes of operators, whose communicative function and cognitive complexity correlates with their scope.[67] Ignoring modality (mostly) and keeping to English examples, the first class ($\pi 1$) are aspectual markers, such as the progressive, continuative (*keep -ing*), prospective (*gonna*, *be about to*), the perfect, and resultative (*be -ed* or irregular form, i.e. *it's broke(n)*). The second class ($\pi 1$) consists of the tense system and habitual/frequentive markers. The third, widest scope operators ($\pi 3$), are proposition oriented modality markers. The ($\pi 2$) class contains event-oriented modality operators, which may lexically be the same as some in ($\pi 3$), but their function is quite different.

The Scope Hierarchy is then ($\pi 1$)-operators $\subset$ ($\pi 2$)-operators $\subset$ ($\pi 3$)-operators, where $\subset$ means less marked. The implication then says that if a language has instances of a more marked class, it also has instances of a lesser marked class. For example there could be languages with only ($\pi 1$)-operators, but not languages

---

[67]This is steeped in the functional grammar of (Dik, 1997).

with only (π2)-operators (π3)-operators. This is confirmed for a wide variety of languages and also investigated in terms hypotheses about diachronic change, usage frequency and other phenomena.

However, it is the acquisition predictions that are of most interest here. Without going into the various permutations of hypotheses, the main prediction is that acquisition of TMA systems follows the above hierarchy in that at least some (π1)-operators are acquired before (π2)-operators, etc. This is confirmed in a wide cross-linguistic survey – including a longitudinal CHILDES corpus study for English. Notice that the (π1)-operators are exactly the early uses of *-ing* and *-ed* discussed early in this chapter, using examples form (Behrens, 2001) and (Wagner, 2001) (e.g. *broke* and *playing* rather than *breaking* and *played*). This gave rise to various hypotheses such as Weist's temporality theory, lexical Aspect first, Aspect before tense, and Prototype theory which are all examined and found to be wanting.

While it may be prototypical to use the progressive with present and ongoing activities, and the simple past (*-ed*) for past and telic events,[68] none of the above theories is seen as an adequate explanation. Instead, she proposes the Discourse Topic Hypothesis (p. 357), which accounts for the initial exclusion of marked combinations, not for grammatical reasons, but for developmental and discourse reasons.

At first, children live in, and only wish to communicate about the here-and-now. Adult speech to children will then be tailored toward this communicative need. For example, talk will be of activities going on in the present and that are perceivable. Importantly, the notion of here-and-now does not apply just to the present moment. Adapted from notions of (Christensen, 2003) and (Klein, 1994), the notion of Topic Time (TT) is utilised. This is like Reference Time, but slightly broader. Applied to the present, TT would include the immediate future and the immediate past. These two (via the prospective and resultative) are included in the (π1)-operators, along with *-ing*.

For a concrete example, compare two situations – a parent doing the washing-up with their child, or the same activity conducted by my flatmate and I. In the first situation, the talk would be of what is happening (*the water's running, now we're putting them in the cupboard*), discrete events that have just happened (*oops, a glass broke!*), or what comes immediately after (*we're gonna have dessert, next!*). In contrast, my flatmate and I would probably talk about what went on that day, certainly anything not as mundane as dish-washing. Instead, the imperfective past might be used, e.g. *When I was walking in town, I met Mary – she says, "Hi."* Discourse confined to the extended here-and-now accounts for a child's initial restriction to the unmarked combinations. Gradually, as there

---

[68]Boland claims that the unmarked combinations, however, are just this – states and activities in the imperfective, telic events in the perfective. There is no space to enter into a discussion about this notion of markedness or the communicative functions and complexity of the various combinations.

sense of time and memory develops, talk will no longer be confined to extended present, and the discourse topics with adults will extend to more remote tense constructions and marked aspectual combinations.

In an echo of what was written in Section 1 of this chapter, Boland writes (p.484) a summary of what a child needs to learn to enter the tense and aspectual system, which is then accounted for by the Discourse Topic Hypothesis along with other mechanisms.[69]

> What is it, then, that children need to learn? They need to learn:
>
> - that parts of event structures need to be selected for building up an adequate description of the property or relation ascribed to the arguments;
> - that there are different event structures;
> - that TT may shift;
> - that event time needs to be encoded; and
> - that TT sometimes needs to be specified linguistically.

While the child begins with an extended present, slowly, broader notions, working in tandem with the child's development along with adult communication about more distant time periods and more complex event combinations. However, the final section of this chapter will focus on the first two bullet points, that of the acquisition of event structure. It yields some rather surprising results, especially the discussion of (Wagner & Carey, 2003).[70]

3.7 Event and Object Individuation

In the first part of this chapter, we examined the analogy between objects and events in the areas of both their perception and individuation. The main factors are spatio-temporal differentiation and conceptual hierarchy. With the former, both events and objects can be individuated by discrete spatio-temporal changes – i.e. the contours of an object (perhaps a rock lying on the ground) or sudden changes in the perceptual stream of an event (say a quick handclap).

But we also saw that both objects and events can be conceptually organised and individuated based on partonomic and taxononomic hierarchies. Unlike the spatio-temporal differentiation, which is largely automatically processed and contains little in the way of conceptual information, these hierarchies are largely

---

[69]Including an appeal to (Slobin, 1985) for event structure acquisition.

[70]Unfortunately (Boland, 2006) only refers to Wagner's early experiments, and none that have to do with intention or what follows in the next section. It is a pity, as combining the results of the two authors raises many questions.

conventional – at least particular instances of the hierarchies are.[71] Moreover, with events we saw the importance of goals and plans in determining and dividing event schemata.

Finally, we looked at experiments conducted by (Zacks et al., 2001) that demonstrated the connection between linguistic description, event individuation, and the interplay between spatio-temporal differentiation (the fine-grained condition) and goal-based differentiation (coarse-grained condition) of everyday, simple events. But what we haven't looked at is how these schemata and taxonomic and partonomic hierarchies are developed – that is, how do young children perceive of events and objects, and do they individuate them on the same criteria as adults? Also, to what extent does language about objects and events effect this individuation? These questions have been studied by (Wagner & Carey, 2003), with some surprising results that could ultimately also cause a rethink of the tense/aspect acquisition studies as well.

3.7.1   Experiment 1 – Object and Event individuation

This first experiment replicates previous work of (Shipley & Shepperson, 1990) on children's object individuation, and adds a parallel condition where children are given the task of individuating events. For the object condition, the subjects are presented with pictures on a computer screen of whole and split objects of the same kind. For example, the screen could contain a picture of several cars divided in half by a small horizontal gap. Other objects used were forks, cupcakes, donuts and trees. Some of the pictures would have only split objects and some would be a mixture of whole and split objects.

Children were then told that they were playing a counting game, and would be shown a series of pictures (as described above) and then with each picture were asked a question that *linguistically* would perhaps influence the counting (object individuation). The options were the general, "How many things are here?" or the kind term "How many X's (e.g. forks or cars) are here?" The split object pictures have the object split, but the two halves aligned and quite close together. Thus, when presented with the kind term, the "correct" answer would be to count the two halves of the fork or the car as one fork or car respectively. That is, if children are indeed using their taxonomic knowledge to differentiate objects. The amorphous "things" would allow for individuation based on spatio-temporal criteria.

The event condition involved showing the children short animations of a variety of simple events. There were events of creation such as building a house, destruction such as a vase breaking, movement of location such as a bunny hop-

---

[71]Knowing that a car is a kind of transport vehicle, or ordering food is part of a restaurant script is certainly conventional.

ping into a hole, or change of state of an object, such as a door being closed. In all cases the events could be conceived as goal based, and importantly they all included spatio-temporally discrete instances of the activity portion. For example, the bunny may take three hops to get into the hole, and the girl may lick the ice-cream cone a few times (with a short break in between) before it is consumed.

Linguistically, both telic and atelic questions were asked in the experiment. A telic question could be "How many times was the house built" with the corresponding atelic question being "How many times did the girl work?"[72] Again, if the children are able to use goal-based individuation for events their responses to the telic questions should be "one time", while they should use spatio-temporal differentiation to individuate the event when given an atelic question (say the number of hops licks of the ice-cream cone). Finally, note that both the object and event conditions were also conducted on adult controls.

Results

In line with the previous study by Shipley and Shepperson, children completely ignored the linguistic descriptions in the object conditions and consistently individuated objects according to spatio-temporal criteria. Thus, when presented a picture of a fork split in two and asked the kind question, "How many forks are there?" the answer would overwhelmingly be "two" (only the five-year olds showed any signs of kind-differentiation). Curiously (and contra the previous study), adults performed exactly the opposite, and consistently refused to make recourse of spatio-temporal differentiation whatsoever. So, when presented with the same pictures and asked the question "How many things are there", the adults would still respond with "one", even though, of course, two half forks may certainly be seen as two different things.[73]

The results of the event trials were a bit more mixed. The adults, as expected, consistently use goal-based counting when presented with the telic questions and spatio-temporal counting when given the atelic questions.[74] But the aggregate data for the 3 to 5 year olds, it is far more mixed. Overall, the children would use goal based differentiation about half of the time when presented with the telic

---

[72]The questions on the atelic activities can be rather difficult. Typically, activities can have a gappy temporal profile, that is, one could be said to have worked form nine to five thought breaks were hopefully taken. A question such as, *How many times did the girl work?* could be rather confusing with no single correct answer, unless the animations somehow non-linguistically made clear that one episode of work was *finished*.

[73]This may be an effect of the experimental setting. Give that the two half objects were aligned and only separated by a small gap, adults may have mentally filled the gap in, so to speak. Still, children got the "thing" condition correct and seem unaffected by the setup.

[74]Appropos of my concerns before, the eat and swim events were still goal based half of the time when given the atelic question, i.e. the three licks of the ice-cream cone were considered one activity of 'eating', while the hopping event was always spatio-temporally differentiated when the atelic question was asked.

description, showing that while the children are still biased by spatio-temporal criteria, their goal-based reasoning is beginning to develop as well. In fact, the results seem strong enough to conclude that children, in contrast to experiments by van Hout, do have a pretty good grasp on telicity.[75]

Looking at the data in more detail, a few anomalies do emerge. While the aggregate figure for the children is fifty percent, the results for the individual descriptions used varied far more. For instance, with the example of "paint a flower", children succeeded eighty percent of the time in using goal-based individuation. This contrasts with the swimming across the river and eating an ice cream cone examples, where the success rates were 28 and 39 percent respectively. Unfortunately, there is no great pattern to this, as the successes and failures do not break down along semantic lines. That is, there was no correlation with the event being a movement, creation, destruction or change of state event. All of the four types showed up both as goal-based and spatio-temporal based differentiation.[76] Instead, it seems to be much more idiosyncratic as to what type of differentiation the children used.[77]

So, it seems that objects and events may not be so parallel after all – at least as far as children's individuation goes. With objects, children ignored linguistic descriptions outright and consistently applied spatio-temporal criteria to differentiate, while adults did the exact opposite. Contrastly, the event condition showed that children could abandon the spatio-temporal strategy when confronted with a telic linguistic description, and instead use a goal-based individuation strategy. Now given these two results, Wagner and Carey wonder whether it is the object test itself that is somehow causing the anomalies.

First of all, there is the difficulty of the ontological status of the split objects – is half a fork still to be considered a fork? If not, then what is one supposed to call it? It is a bit optimistic to expect three year old children to grasp these deep matters, when they are still the matter of great debate in late night university undergraduate drinking sessions. Moreover, there is also one rather significant dis-analogy between the object and event conditions. With the objects, the alternative to the informative kind term was the fairly uninformative "thing" that was used for all of the objects.

But for the event trials, the atelic description was still something more informative like "work" or "hop", that are commonly named and different activities. Therefore, they conducted a second experiment on object individuation that involved not just split objects, but also the contrast between collections and individual objects (forest vs. tree) and part/whole (wheel on a car). This has not only the advantage of making both descriptions contentful, but also brings into play children's knowledge of partonomic hierarchies, making the parallel to

---

[75](Hout, 1998) only used examples with the notoriously difficult eating and drinking predicates, which even the adults in the current study had difficulty with.

[76]There was also no difference based on the syntactic form of the description.

[77]Perhaps it has to do with how many of the event schemata the children have learned.

events much closer.

3.7.2 Experiment 2

Experiment two was conducted on (different) sets of three and five year olds, and consisted of three distinct conditions. The first condition is merely the split object test of the first experiment, while the second two were designed to further test the hypothesis of Shipley and Shepperson that young children only count spatio-temporally determined objects, no matter how (linguistically) they are asked to count.

The first of these involved a contrast between collection and object. The collective terms used were things likes "forest" and "family". A picture might be two separated forests consisting of four trees each (where the trees are grouped closely together and are slightly overlapping). The desired answers to the questions "How many forests?" and "How many trees?" would be "2" and "4" respectively. Note that this experiment is slightly easier than a similar one done by Shipley and Shepperson's where children would be shown, say, a picture of three dogs and three cats and asked to count both the animals and the *kinds* of animals. In both cases children counted individuals, but this is attributed to the fact that understanding the logic of super-ordinate classification is difficult for young children. Rather, in this condition, the contrast is a bit simpler – comparing individuals to a collection – based on the linguistic description.

The second of the new conditions involves a part/whole contrast. The stimuli were pictures of ordinary objects with their commonly named parts, e.g. a bike and its wheels or a butterfly and its wings. Thankfully (especially in the butterfly case) the parts were still attached to the objects and children would be asked to count either the amount of butterflies or the amount of wings. If linguistic description is ignored by children, then the answer to both queries would be to count the whole objects. However, it is well known that young children learn body parts amongst their earliest words and can readily learn novel part names, provided they already know the name of the object. Moreover, while the parts may indeed still be attached to the object, a case can be made that the parts (wheels and wings) have a distinct spatial contour in relation to the rest of the object, and that spatio-temporal differentiation would serve in this case as well. Therefore, in this condition at least, it was predicted that even 3-year-olds would be quite successful.

The results confirmed exactly this – in the part/whole trial, 5-year-old children performed at adult levels, while the 3-year-olds weren't too far behind. The results of the split-object condition were replicated from the first experiment, with children counting the half objects no matter what description they were given. So far, this is just what was expected; however, the results for the individual/collection condition are rather interesting. Statistically, they are almost an

exact match of the event individuation condition in the first experiment.[78]

So, it does seem that contra Shipley and Shepperson, linguistic descriptions can indeed influence even 3-year-olds object individuation, and that counting for them is not confined to counting spatio-temporally defined whole objects. If asked to count wings, even the youngest children would rarely count the butterflies the wings are attached to. Also, if asked to count families, at least half the time, the youngest children did so, instead of counting the spatio-temporally distinct family members.

Nevertheless, this is not to say that children behave in a completely adultlike way in regard to both object and event individuation. The split-object condition shows that in difficult cases, both children and adults will pay less attention to linguistic description and use their preferred individuation strategy – spatio-temporal in the case of children, and kind-based in the case of adults. Indeed, the results do indicate that while children do pay attention to linguistic descriptions, there is still a rather robust spatio-temporal bias.

Building on work of (Carey & Xu, 2001), they propose the existence of an object tracking system that appears quite early on in a child's ontogeny – that is objects are tracked and individuated on the basis of spatio-temporal information. While infants also can infer goals when presented with individual events (we saw this quite early in the chapter), it is probably the case that they have yet to fully integrate this talent into developing the necessary hierarchies and schemata that are coded by language. Thus we get the kind of mixed results seen in the event individuation condition and the individual/collection condition.

Wagner and Carey take pains to point out that this tracking system is not to be considered a step on the road to more complex kinds of individuation. For example, kinds cannot be spatio-temporally defined, so individuation in these cases must come from knowledge of taxonomic hierarchies. Moreover, if adults were presented with a picture of conceptually unrelated objects, their only recourse to counting them would have to be done on spatio-temporal criteria.

However, in the context of language acquisition and learning in general, I think a case can be made for a process happening. There are many ways to differentiate objects, and these ways can be seen as different perspectives – in the same way that different tense and aspect combinations can throw a different viewpoint on events. Here, Wagner's results show the process involves beginning with a default individuation strategy – spatio-temporal tracking. However, as the child learns language and develops conceptually, other possibilities of perspective arise. But while even the 3-year olds may be sensitive to linguistic cues, they are not yet able to switch between different individuation criteria at adult-like levels. Learning the criteria requires a wealth of world knowledge and experience and a linguistic sophistication to pick up subtle markings (count/mass, telicity).

---

[78]Although there is no mention as to whether the data breaks down on individual pictures in the same way as the event individuation condition.

It is then unsurprising that when confronted with a task that is a bit too hard, children will revert to what they do best – spatio-temporal tracking.

3.8 Conclusion

We have travelled through a rather long journey through psycho-linguistics – first in cognitively grounding the concept of 'event', and then seeing how children learn to understand them and talk about them in the context of tense and aspectual uses. While no simple conclusion can be drawn from this excursion, there are several lessons to be learned.

3.8.1 The learning process

First, however the mapping process works, whether it is more generative in nature, usage based or something in between, both the acquisition of the necessary concepts and applying these to a specific language is a lengthy developmental process. Where, as (Weist et al., 1997) elegantly points out, the relation between the acquisition of a specific language and the development of concepts most likely have feedback effects upon each other. For example, in comprehension, Polish children had an advantage in the temporal dimension, while the Finnish children had a head-start in the spatial dimension. These discrepancies are put down to the relation between the languages' systems for expressing spatial and temporal concepts and there accessibility to information processing strategies.

More specifically, his neo-piagetian developmental timeline is a nice illustration of the foundations required for tense comprehension. The child begins with a fairly impoverished notion of the relationship between events and time (speech time only) and within a few years develops a full-blooded temporal conception, where she can not only locate events in time, but can coordinate three separate temporal intervals. As seen in the introduction, these two ingredients are exactly what is used in the event calculus in order to interpret such tense constructions as the past perfect, or deal with adverbials such as *before* and *after*.

Interestingly, the experiments in (Weist et al., 1991) highlight both an important methodological point. In the production tests Finnish children used the perfect tenses earlier than their American or Polish counterparts. At first glance it may seem to be the case that they have developed the Free Reference Time system earlier. However, the comprehension tests show that cross-linguistically the transition to a Free Reference Time system begins around 4;6 and is not mastered (on Weist's tests) until 5;6 and 6;6. The explanation for the Finnish children succeeding earlier on the production tests is that there linguistic environment contains more instances of the perfect tenses than in the environments of American or Polish children. This does tell us that mere use of a morphological form does not indicate that a child has any sort of mastery of it. While there may

be influences from language on thought and vice versa, in this case a child needs to be able to have enough working memory to keep track of the various goals and the computational power to co-ordinate them properly, before they can be said to comprehend something like the past perfect.[79]

The experiments done by Wagner shed some interesting light and a critique of Weist's experiments. In (Wagner, 2001), she first replicates results done by Weist, that show that even 2 year olds can understand simple tense.[80] But in a follow up experiment, she varies the parameters enough to make an issue of whether an event is completed or incomplete. The results show that the two-year-olds are successful on the test only when the past tense scenario was a completed event and the present tense event is (still) incomplete. This result is presumed to be a qualified endorsement of the Aspect First hypothesis, that is children are using tense morphology[81] to initially be an indication of completion information (past tense means complete, etc.), and is only later in the child's development used to mark actually temporal relations.

Caveats

But,[82] there may be a different explanation for the results in (Wagner, 2001). (Wagner, 2002) shows that when deprived of agentive information in the scenarios of the experiments, the children do not succeed on the aspect tests until the age of five.[83] From this experiment she concludes that possibly it isn't until around age 5 that aspect's completion entailments are understood. Prior to this, children are relying on agentive information and intention to make their judgements. It may then be the case that rather than the competing theories of Tense First or Aspect First, it may really be *Intention First* that is the child's initial strategy, with a full understanding of the tense and aspectual distinctions in a given language completing some years later.

Examine this in light of the work on events proper summarised in (Zacks & Tversky, 2001). Even understanding what an event consists in requires on levels of complexity more than a simple spatio-temporal change, it also requires knowledge of intention or even conventional scripts that can only be learned by living in a particular culture, and only made sense of if one understands the incredibly complex web of shared intentionality that we are all part of. The

---

[79]Nevermind adding the causal dimension seen in the mini-discourse examined in the introduction.

[80]Methodologically she makes an important refinement by comparing all three tenses, rather than a past/non-past distinction that could reduce to the ability to make a realis/irrealis distinction.

[81]As this experiment was conducted only on English speaking children and only with the present/past progressive. So, tense morphology here boils down to *is* and *was*.

[82]This is not mentioned by Wagner herself.

[83]Note that this meaning of aspect is rather impoverished and conflates perfective and imperfective with completed and incomplete, which isn't quite the case.

event calculus takes not only the complexity of integrating events for tense and adverbial combinations rather serious, but also the notion of intention and goal.

At the heart of the representation of the accomplishment-style aktionsart in the event calculus are these notions. An activity, can if successful, lead to a goal.[84] There is the possibility of layers of hierarchical organisation where a sub-event is a goal that looked at from a wider perspective, this goal can then be recruited to be a sub-event of a larger event with a different goal. This is a direct parallel to the notion of script in (Zacks & Tversky, 2001) and its partonomic structure.[85]

Also, as seen earlier, computing tense structures also involves in having goals of successfully locating the event and integrating the event in context. Here, there is no real room for Aspect or Tense first, but both are needed.[86] This is obvious when one needs to interpret, say, a past, progressive sentence, but even to interpret a simple, past tense sentence is not merely locating an event in the past, but also to embed it in the proper discourse context. That is, states or processes (both of which are fluents in the event calculus) are the bedrock upon which events rests.

A broader outlook

To finally return to the acquisition experiments, (Wagner & Carey, 2003) throws a spanner in the works. The lesson to be drawn is that children and adults do not conceptualise events in quite the same way. Looking back at the psycho-linguistics of events, this is really no surprise. While knowledge of intention, hierarchical and taxonomic thinking are required,[87] the various scripts for events are learnt,[88] and even whether a break in an event is considered an actual termination or just a pause to be ignored can depend upon the specific event or the context.

But, it does cast doubt on the previous tense and aspect experiments in general. If children conceive of events slightly differently, interpreting their answers according to an adult paradigm may perturb the statistical results. Getting an answer or a picture match 'wrong' may not be a tense or aspect mapping immaturity, but merely a misunderstanding of what kind of event is talked about. As an open question, I do wonder exactly how this conceptual difference cashes itself out. Is the misconception a matter of cultural knowledge, or does hierarchical

---

[84]Often when there is a sentient agent performing such an activity, we can say that the goal is there intention. By extension, a ball rolling inevitably of a cliff, can also be see as an activity leading to a 'goal', though without a human intentional element.

[85]It also sees favourable parallels in the notion of rational action in (Gergely & Csibra, 1993).

[86]Along with a proper conception of the various Aktionsart types.

[87]This can be seen as the building blocks of an event script. The specifics may be entirely conventional.

[88]That is not to say that all people in a given culture will have precisely the same notions of any given event script. I believe the differences are often at the heart of the common English expression that comes out in an argument when both sides realise there's no convincing the other – *That's just semantics!*.

planning and the ability to have narrow or broad perspectives on a given situation also emerge slowly, causing further interactions with the learning of tense and aspect, both conceptually and linguistically?

However, the problems brought up by (Wagner & Carey, 2003) should not cause us to despair altogether. While it does have an important lesson to teach us in how to properly conduct and evaluate acquisition experiments, the entire view of the acquisition process and event conceptualisation need not be pushed to one side. While the experiment points to the strong possibility that children's event scripts are not quite the same as those of adults. There is no reason to things that the 'building blocks' of event conceptualisation are any different. This experiment finds children focussing primarily on the spatio-temporal view of events, while (Wagner, 2002) shows the reliance of young children upon the agent's intention to determine whether or not an event has been completed.

Both of these are basic ingredients in (Zacks et al., 2001) definition of event as well as how events are constructed in the event calculus. Any temporally extended event includes a process that can be identified with a plan or intention that leads toward the event's goal.[89] The event calculus has an ontology of events and processes along with a general planning and causal mechanism that relates them to each other. Wagner's experiments only serve to show that this is exactly what children are using in conceiving of events, determining completion conditions, and eventually locating them in time and in respect to each other.

What children may not have is the adult's event 'lexicon', i.e. the conventional scripts for everyday events are not only not consistent across cultures, but as the coercion phenomena shows, can even vary for a single speaker depending upon context. Moreover, when dealing with events whose definition is quite dependant upon cultural values, there is no reason to expect such things to be universal across cultures or completely comprehended by three year olds, who still have many years of living within a culture before becoming an 'expert' on it. Ultimately, we should think of what is 'universal' as things such as basic cognitive mechanisms used to structure time, create event schemata, and various temporal and intentional perspectives.

For example, what is a 'stealing' event? While there may be stereotypical connotations of a man in a mask with a gun and a swag bag, the push of a button by a devious banker can be stealing all the same. Ultimately, for the more 'higher level' events, the agent's intention is not enough, but rather the intentions and values of an entire society are needed to fully comprehend it.[90] This takes time

---

[89]Intention should be taken here in the broadest sense. It can be applied to my intention of walking to the store or a ball's 'intention' of rolling into the street. In the latter case, the spatio-temporal profile, along with naive physics takes the place of agentive intention.

[90]A baby in their own baptism (pretend it happens at 18 months instead of in infancy), will perceive a man in strange robes talking, another adult agreeing to renounce Satan and water being poured on his head. These are (some of) the spatio-temporal parts of a baptism. But, without an understanding of the role of religion in a culture, this particular acts are really quite

and life experience, and aside from a move to innate concepts, there really is no substitute.
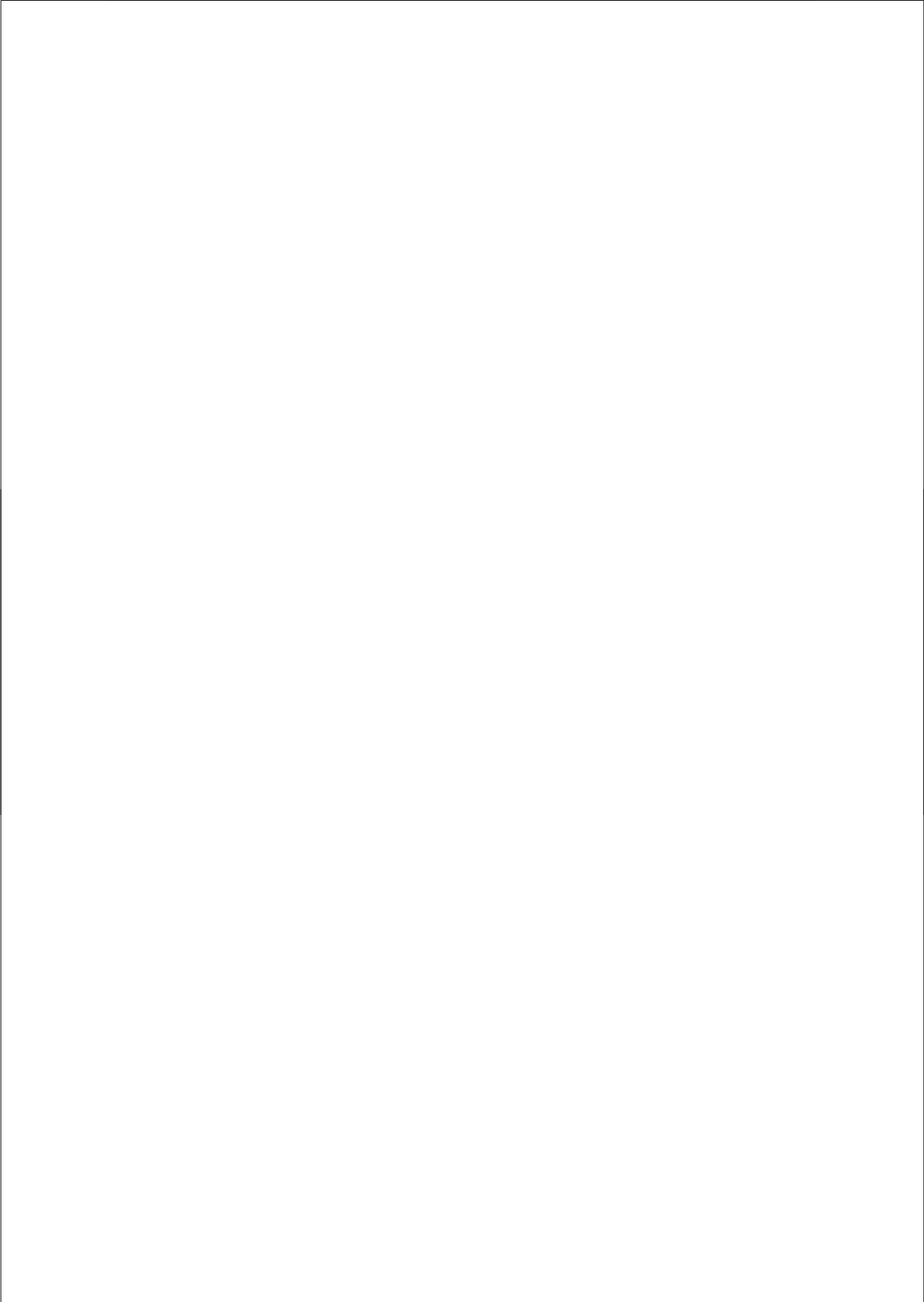
3.8.2   Looking Forward

We will look forward by looking back at the introduction to this chapter. We began by looking at, intuitively what it takes a hearer to interpret a simple two-sentence discourse. Ostensibly simple, this process is actually incredibly complex. This was then reinforced by looking at the work of Weist, who postulates a development of temporal concepts and also gradually developing more sophisticated ways of relating the events. In the interlude, we will see the event calculus formally and how its formal mechanisms parallel the RRT stage of Weist, as well as how the informally described discourse interpretation process can be stated formally. In a sense this can be seen as the 'universals' involved in tense and aspect production and comprehension. But, following the work of (Zacks et al., 2001) (amongst others), the schemata developed, while utilising such generalities as planning, intention, and spatio-temporal change are a default 'lexicon' that can change in an ad hoc way (due to a specific contextual anomaly), but can vary in more patterned ways within and across cultures.

The following interlude presents an informal introduction to the Event Calculus, with special emphasis on the way events and their mapping to verbs and VP's is modelled. It is no accident that the notion of event schemata and hierarchical structure and the role of intention and goal play a prominent role. Moreover, the results of the last experiment of Wagner should not be seen as a reason to despair. Combining her results with the Zacks and Tversky notion of event structure in mind, the event schemata in the Event Calculus use the universal building blocks such as spatio-temporal change, object tracking, and goal directed orientation. However, as the resulting schemata of any VP,[91] is considered a default and often contextually determined, the possibility that children do not have quite the same specific conceptions as adults no problem, or surprise, whatsoever.

---

meaningless for him at the time.

[91]These are called scenarios in the Event Calculus.

Interlude: An introduction to the Event Calculus

In[92] the previous two chapters, the notions of Aktionsarten, aspect and tense were discussed, followed by the ways these are manifest psycholinguistically. For the latter, the notions of event schema, goals, plans and intentions played a paramount role in the construction of various events. Furthermore, it was also seen that when one is relating multiple events to each other, viewing them through various aspectual construals, or anchoring them in time relative to a discourse context, these very same notions crop up again. Before moving on to the latter three chapters that are studies of specific semantic and linguistic topics that will involve formalisation, a brief introduction to the formalisation used is in order.

The Event Calculus (Lambalgen & Hamm, 2005), is a formal semantics for tense and aspect that takes inspiration form a planning formalism from robotics. As human beings are not robots, this may sound a bit disconcerting at first, but as a formalism for computing tense and aspect, the ingredients of this sort of planning fit perfectly. For example, suppose a there is a goal specified for the robot to pick up a package in a building. Starting from this goal, a plan is then computed, which consists of a sequence of actions that will get the robot to the package's location. This sequence is obtained by backward chaining from the initial goal through a sequence of sub-goals ending with the robot's current state and position.[93]

This sort of computation requires a model of the world, which is in this case a map of the building, a causal theory of the robot's actions, various variables like 'door open/closed', its initial position, and a record of its actions. Furthermore, the robot will need a set of actions it can perform (*walk down the corridor, turn right, go through the door*), as well as possible observations it can make, such as whether the door it needs to go through is open or closed. With this, a plan

---

[92]Much of this exposition borrowed from (Hamm, Kamp, & Lambalgen, 2006).

[93]Recall the notion of goal, plan and subgoal in storytelling and comprehension from (Trabasso & Stein, 1994).

to pick up the package can then be computed. While the robot is executing the plan, its world model is updated with the various observations and actions it makes, and the plan may need to be recomputed on the fly, should something about the world model differ from the original plan. For example, a door could be closed that was open in the original plan's world model, or worse, the robot makes the observation that a corridor on its original building map is now inaccessible. Finally, the actions of the plan include can be either instantaneous or continuous such that the former can take place during the latter. For example, there could be the action of moving along the corridor at a certain speed, and during this extended activity, there could be various instantaneous activities such as taking a observational reading of the current distance between the robot and the wall.

These notions can be harnessed for the computation of tense and aspect. Upon hearing a new sentence, a listener has an initial discourse model and must integrate the sentence into the initial discourse model. If the sentence is a tensed non-stative sentence, the actions and participants are unpacked and the discourse model is updated with these. For example, if I hear *John is running in the zoo*, I need to update my discourse model with the action and the agent of the action.[94] It could be more complicated, however. For example, if I hear *John is going to feed himself to lions*, it is not only an action like *running* that is expressed, but also the existence of a plan directed toward the goal of non-existence. Should the sentence be a simple stative, like *John is suicidal*, this is an analogue of the robot's more mundane observations about doors and distance, and my discourse is updated with a property about John's mental health.

The appendix contains a more formal description of how this works and the reader is encouraged to see (Lambalgen & Hamm, 2005) for a much more complete and elegant formal picture. However, in the following sections of this interlude, we shall give an informal introduction to the language and workings of the event calculus, in order that the reader will be able, at least, to have a grasp of the more formal sections of the later chapters.

3.9   Language and ontology of the Event Calculus

3.9.1   Basic Ontology

As the above discussion implies, the event calculus is planning formalism, whose language talks about actions, goals and causal relations. Given a goal, initial state and casual relations, it returns a plan to achieve this goal. It is a many sorted logic, that has two basic sorts of events, for viewing events either perfectively or imperfectively. The first are event types, symbolised as $e, e', \ldots, e_0, \ldots$. The

---

[94]And the time that it is happening.

second are called fluents, symbolised as $f, f', \ldots, f_0, \ldots$. Event types are various kinds of actions, such as *fall* or *break*. Fluents are time-dependent properties such as *being broken* or *walking*. While the time parameter of fluents is implicit, they can have further parameters, e.g. for the subject of *walking*, which come into play, especially when examining continuous change and incremental themes.[95]

As far as the ontology is concerned, there are also sorts for individuals (e.g. 'John'), real numbers that can be interpreted as instants of time, or other quantities for position, velocity or the degree of a quality.[96] Importantly, this sketch of ontology is concerned with discourse models and not directly with the real world. For example, as seen in Chapter 2, there are no perfective and imperfective events in the world, but are rather different ways of conceptualising the world.

3.9.2 Basic Predicates

Instantaneous Change

The event calculus has a set of predicates for handling two types of causality – instantaneous and continuous change. We will first explicate the predicates need to describe the former, such as the classic example of two balls colliding.

- $Initially(f)$ ('fluent f holds at the beginning of the discourse')
- $Happens(e, t)$ ('event type e has a token at t')
- $Initiates(e, f, t)$ ('the causal effect of event type e at time t is the fluent f')
- $Terminates(e, f, t)$ ('the causal effect of event type e at time t is the negation of the fluent f')
- $Clipped(s, f, t)$ (roughly, 'an event type terminating f has a token between times s and t')
- $HoldsAt(f, t)$ (fluent f is true at t)

---

[95] Specific events and fluents are derived by *reification*. For example, as in standard logic translation, the verb *walk* is represented by a predicate $Walk(\overline{x}, t)$ (all free variables exhibited; t is the temporal parameter). While we shan't go into the details here (see (Lambalgen & Hamm, 2005, Chapter 6)), the expression $\{t \mid Walk(\overline{x}, t)\}$ (with the $\overline{x}$ as free parameters) may be formed, which we can think of as a fluent $walk(\overline{x})$ which contains an implicit temporal parameter. In practise, we will usually leave out the free parameters where there is no possibility of confusion, and refer to the fluent as *walk*. In general, the fluent $f(\overline{x})$ can then be an argument of the $HoldsAt$ predicate (see next section), e.g. $HoldsAt(f(\overline{x}), s)$, intuitively meaning $s \in \{t \mid A(\overline{x}, t)\}$, i.e. $A(\overline{x}, s)$. Events types $e(\overline{x})$ can be constructed from a (Gödel number of) a formula $\exists t A(\overline{x}, t)$, which abstracts away from the temporal parameter.

[96] While real numbers are used to represent instants of time in the formalism, the actual mental construction of time is much more complicated. Concepts of order duration and temporal perspective involve a construction goals to be obtained, and plans to achieve them (e.g. for future time). The past (or remembering it, at least) can be constructed analogically. For our purposes, we will only concern ourselves with real numbers as instants of time, but the reader is invited to see (Lambalgen & Hamm, 2005, Chapter 2).

Imagine an example of Ball A lying motionless on a table and being collided with by Ball B, causing Ball A to move. If Ball A then hits the edge of the table, its motion is stopped.[97] So, we then have *Initially*(*still*) to describe Ball A's initial lack of movement. The causal information can be described as *Initiates*(*collision, move, t*) which means that if Ball A is hit by Ball B at time t, the causal effect is the fluent *move* for Ball A. The second causal effect is *Terminates*(*hit-edge,move,t*) which says that the event of hitting the edge of the table negates the *move* fluent.

The predicate *Clipped* can be used to describe that an event token terminating the ball's movement has happened between times s and t.[98] For example, imagine that our scenario begins at t=0, with the Ball B collision happening at t=2 and the edge collision happening at t=6. Assume for now that the ball is moving between times 2 and 6.[99] For example, we would have *HoldsAt*(*move*, 5). But since our scenario contains *Happens*(*hit-edge,6*), we know from *Terminates*(*hit-edge,move,t*), that movement is stopped. We can use the *Clipped* predicate to express this, e.g. *Clipped*(3, *move*, 8), which says that between times 3 and 8 a token of an event-type terminating *move* happened, in this case *hit-edge*.

Continuous Change

The above predicates for instantaneous changes nicely describes what happens when billiards balls hit each other,[100] but is inadequate to describe say the process of filling a bucket of water from a tap, exactly what we need to account for activities or accomplishments. For continuous change, we need two more predicates.

- *Trajectory*(f_1, t, f_2, d)
- *Releases*(*e, f, t*)

To explain these, let us continue with our bucket-filling example. Let the fluent *fill* be the activity of water filling the bucket, and the fluent *bucket*(*x*) be the partial object that represents the height of the water in the bucket. Suppose that we begin with an empty bucket (e.g *Initially*(*bucket*(0))).[101] Imagine there

---

[97]This would make for a very boring game of pool, but it makes the explication of the predicates much easier.

[98]It is a critical predicate for capturing the principle of inertia that will be seen in the next section.

[99]Intuitively, this is the principle of inertia. We know that *Happens*(*collision*, 2), which initiates movement, and have no information that movement has been terminated, and thus can assume that the ball is still moving.

[100]And in Aktionsart terms, is enough to account for achievements and other instantaneous event semantic event types.

[101]This is an example of a parameterized fluent hinted at above. Here the parameter in *bucket*(*x*) refers to the degree of the buckets fullness. Initially it is empty(i.e. 0).

is an event *turn-on* which is the event of turning on the tap. This event initiates *fill* which is the process of water flowing out of the tap, and into the bucket. Intuitively, as long as the water is flowing, the bucket is being filled to a greater degree.[102] This is exactly what the *Trajectory* predicate is used for. For example, the predicate

$$Trajectory(fill, t, bucket(x + d), d)$$

says that if $fill$ is true between times t and $t + d$, then at time $t + d$,[103] the bucket is filled to the degree $x + d$. We can also nominate a constant c to represent the degree when the bucket is considered to be completely filled, and thus ending the filling process.

The *Releases* predicate is necessary to bring together the two notions of change. In this example, we have $Initially(bucket(0))$, which also is the same as saying that $HoldsAt(bucket(0), 0)$.[104] We have seen how events can trigger instantaneous change, e.g. $Initiates(turn\text{-}on, fill, t)$, and that $fill$ gradually changes the state of the bucket, but not the relation between the $turn - on$ event and the degree change in the bucket. This is accomplished by

$$Releases(turn\text{-}on, bucket(x), t)$$

which ensures that once the tap is turned on, the state of the bucket is free to change. This ties in with the *Clipped* predicate described above. As $bucket(0)$ is a fluent, it needs to be terminated in order for the continuous change to take place. One way of something being *Clipped* is for an event to happen that releases it.

3.10 Aktionsarten and Scenarios

In the previous section we discussed the predicates of the event calculus as well as events and fluents and how the predicates relate them to each other. The examples given to illustrate them did look like examples of the various event types discussed in Chapter 2. This section will make the basic Aktionsarten more explicit, as well as introduce the notion of scenario.

First of all, we need to slightly differentiate the types of fluents. In the last section fluents were used to represent both states such as the non-movement of a ball, or the degree of fullness of a bucket. They were also associated with activities such as the movement of the ball or water filling the bucket. While all fluents are

---

[102]Until the bucket is completely filled. After this, the water may still be running, but we would construe $fill$ as having stopped.

[103]Using the predicates, this means that $HoldsAt(fill, t)$ and $\neg Clipped(t, fill, t + d)$.

[104]This follows from an axiom (see Appendix) that says that if a fluent is initially true in the discourse model, then it holds at time 0.

constructed by the same reification procedure, there can be differentiation via the syntactic role they play in the scenario. This can be seen below with the general notion of eventuality that the basic event-types are constructed from.[105]

> An *eventuality* is a structure (f_1, f_2, e, f_3), where
>
> 1. f_1 is a fluent which represents an activity, something which exerts a force
>
> 2. f_2 is a parametrized fluent, representing a parametrized object or state, which is driven by the force f_1
>
> 3. e is the culminating event, representing a canonical goal
>
> 4. f_3 is a fluent which represents the state of having achieved the goal
>
> Accordingly, one may associate to each VP a quadruple, each element of which is of the form '-' (indicating that this slot may remain empty), 'e' (third argument only) or 'f' (first, second and fourth argument)

This is very similar to the 'event nucleus' of (Moens & Steedman, 1988) discussed in Chapter 2. Here f_1 corresponds to the preparation, f_3 to the consequent state and e to the actual event. However, there is also the parameterized object/state, f_2, which we used in the bucket-filling example, and in general is important for the treatment of accomplishments and the progressive. These differences are between state and activity fluents (f_2 and f_3 vs. f_1). The first can appear as the fluent argument of the *Releases* predicate and the second fluent argument of the *Trajectory* predicate, but not the first fluent argument of the predicate. In contrast, f_1 appears as the first argument of the *Trajectory* predicate. Recall the bucket-filling example with $Trajectory(fill, t, bucket(x+d), d)$, where the first fluent is the activity of filling the bucket, while the second fluent argument is the state of the bucket's fullness.

Aktionsarten can then be defined as specific types of eventualities. Note that the above defines an eventuality as having the form (f_1, f_2, e, f_3), but not all slots of the tuple need be filled. For example, the if the tuple is $(-,-,-,f_3)$, this is an Aktionsart that contains only a stative fluent. Unsurprisingly, this is the eventuality for a state. We shall now examine the various types of eventualities and the corresponding lexical and conceptual material, the scenario.

---

[105]Definition stolen from (Lambalgen & Hamm, 2005, p.84). This definition can be seen as the 'bare minimum' to capture the basic types of eventualities. However, the aspectualisers discussed in Chapter 2 require extra structure. For example, an ingressive event would have a punctual event that is prior to the activity fluent.

3.10.1 Scenarios

While leaving the formal definition to the appendix, scenarios can be seen as 'micro-theories' that state specific causal relationships holding between predicates, as well as events that have occurred in a given situation. For example, 'fill a bucket' required a *filling* event and a partial object corresponding to the bucket's state, as well as the events and causal relationships (i.e. the $turn - on$ event and the *Trajectory* predicate) that are involved. In general a scenario is a conjunction of statements, e.g. $S(t) \rightarrow Initiates(e, f, t)$,[106] where for now, we can think of a state as a positive or negative *HoldsAt* predicate.[107] More importantly for our purposes the scenarios express lexical meaning. To see how this works, we shall see examples of achievements, activities and accomplishments.

Achievements

Achievements such as 'reach the top' consist of an instantaneous event followed by a resulting state. Therefore, it will have a form (-,-,+,+) which is the tuple version of the preceding predicate. The relevant fluent and event are constructed by reification, with the lexical entry for 'reach the top' being

$$Initiates(reach, \textit{be-at-the-top}, t)$$

which says the at time t the event *reach* initiates the state of being at the top. Note that the terminating event terminates an activity that is not explicitly mentioned in the scenario. But world knowledge and the context of the discourse suffice to fill this in.

The even simpler eventuality of a point (e.g. *cough*) is represented as (-,-,+,-), meaning that it consists of an instantaneous event only.

Activities

The previous example was of eventualities regarding instantaneous change (achievements) or no change at all (points). In this section we will look at activities. Take the example of 'pushing a cart'. There is continuous change regarding the cart's position that is driven by the activity *push*. Moreover, there are punctual events *start* and *stop* that can be said to initiate and terminate the pushing. With these

---

[106]The predicate *Initiates* is in the consequent, but it could be any of the other predicates of the Event Calculus, barring *HoldsAt*. The antecedent may also be empty.

[107]But equalities between fluent terms or between event terms or formulas of arithmetic are also allowed. Secondly, the $\rightarrow$'s should not be taken to be a material conditional, as the Event Calculus is non-monotonic. However, we shall leave the meaning of the arrows to the final sections of this interlude that discuss the workings of the system. For now, however, keep in mind the notion of backwards planning from a goal discussed at the beginning.

as well as a parameterized f_2 fluent $position(x)$ to represent the cart's positions and a function g to help calculate it, we can get the appropriate scenario.

1. $Initially(position(a))$[108]

2. $Initiates(start, push, t)$

3. $Terminates(stop, push, t)$

4. $Releases(start, position(x), t)$

5. $HoldsAt(position(x),\ t) \rightarrow$
   $Trajectory(push,\ t,\ position(x + g(d)),\ d)$[109]

This scenario states that the cart has an initial position, and once the pushing is started by a starting event, the cart's position is allowed to vary, via *Releases*, and the fact that the pushing causes the cart to vary in position. In general, we will refer to statements of the form of the last two as 'the dynamics'. Note that once the *stop* event occurs,[110] the pushing activity is terminated.

Accomplishments

An accomplishment is similar to an activity in that it involves continuous change, but does differ in the latter is atelic, while the former involves a goal. This can be done by defining a stage of the partial object where the telic goal is reached, meaning that the eventuality is completed and finished. For example, in the accomplishment 'build a house', we have a activity *build* that causes the partial object $house(x)$ to change. At some point, the house is completed, which triggers a $finish$[111] event that terminates the building process – call this constant c, with $house(c)$ the fluent representing the consequent state of the house being built. In general, the eventuality corresponding to an accomplishment is $(+,+,+,+)$, and in the specific case we are looking at, it is $(build, house(x), finish, house(c))$. Its scenario could be

1. $Initially(house(a))$

2. $Initiates(start,\ build,\ t)$

---

[108]a represents whatever the carts position is at the beginning of the scenario.

[109]The g in the *Trajectory* predicate is merely a function of time that is definable in the language of the reals that satisfies $g(0) = 0$. This function means that (in this scenario) that we needn't assume uniform velocity for the pushing, which would be the case if the description of $position(x)$ in the third *Trajectory* place would have been merely $x + d$.

[110]This could be reaching a destination, or merely the pusher getting tired and stopping the pointless exercise.

[111]For activities there is a *stop* event that is not dependent on any condition.

3. *Initiates(finish, house(c), t)*

4. *Terminates(finish, build, t)*

5. *HoldsAt(build, t) ∧ HoldsAt(house(c), t) → Happens(finish, t)*

6. *Releases(start, house(x), t)*

7. *HoldsAt(house(x), t) → Trajectory(build, t, house(x + g(d)), d)*

Note that the only major difference between an accomplishment and activity lies in statements 3 and 5 which involve the goal *finish*. The former says that the *finish* event initiates the state of the house having been completed, and the latter describes the conditions for reaching this goal.

Final Aktionsarten remarks

While we have spoken of the eventuality of a VP, or a scenario for a specific VP like 'build a house', it should be noted that these need to be seen as default correspondences. We can view scenarios as Zacks and Tversky's event scripts made flesh. There it was shown that scripts are not set in stone, but can vary due to context, culture, or even as seen in the final experiment of Wagner, age. For example, the results in (Wagner & Carey, 2003) strongly suggests that young children are not using conventional event schema for events, or at least do not use the same divisions. In the chapter on lexical aspect, we saw that certain verbs were of different aktionsarten in different languages,[112] and it was also noted that such things can vary even between speakers of the same language. What is important is that we have a set of 'building blocks' to create the various scripts in a part-whole manner, that takes the notion of goal and intention seriously.[113] There is also another reason to view scenarios as flexible in their details.

In Chapter 2 numerous examples of 'coercion' were discussed, which means that the mapping of the VP to an eventuality type is also a default at best. For example, there were Comrie's observations on the fine conceptual line between viewing something as a state or an activity (across languages), as well as how some verbs in the same language can be viewed as either.[114] Take the example from (Croft, 1998)

(9)   She is resembling her mother more and more every day.

---

[112]Especially those that can be seen as either states or activites, i.e. *dry*.

[113]In the next chapter, events (in specific contexts) will be seen that differ only in intention and social information.

[114]Recall also that in the event calculus, both states and activities are represented by fluents. It is only their syntactic role in the eventuality that differentiates them.

where the verb *resemble*, normally a state is used as an activity. In terms of the Event Calculus we would need a way to 'transform' what is normally a mere f_3 fluent to an activity.[115] Informally, we can see how this is done. The default mapping yields a fluent *resemble*. The extra lexical material, 'more and more' introduces a parameterized fluent $f'(x)$ that denotes *resemblance to degree x.* Finally, given that the sentence is in the progressive (and therefore an activity in this context), we need to introduce our two dynamic statements

1. *Releases(e, f_2, t)*

2. *HoldsAt$(f_2(x),t)$ $\rightarrow$ Trajectory$(f_1, t, f_2(x'), d)$*

    To begin with there is a bit of a mismatch. Normally, the fluent *resemble*[116] is represented by an f_3 fluent, but the additional lexical material of the sentence demands a fluent for continuous change as well as a dynamics. So in this case, the fluent *resemble* is 'unified'[117] with f_1 of the above scenario, with $f'(x)$ unifying with the f_2 fluent. Thus what is normally a stative eventuality becomes an activity.

3.10.2   Hierarchical Planning

Thus far, the basic elements and predicates of the Event Calculus have been introduced as well as a way of linking lexical and conceptual information to build scenarios for discourse models. However, the reader may have noticed something lacking in our discussion of events proper. Up to now, all the examples of e are instantaneous, for example, points such as *cough*, or *start* and *finish* events used in a scenario. However, something temporally extended, like 'build a house' is certainly an event as well, and we would like in the language of the event calculus to say that it happens at a certain point in time.[118] But, examining the eventuality and scenario for it in the last section, there is no feasible candidate. The only e-types are the *start* and *finish* events, which are punctual.

    Moreover, there is no way to account for the perfective/imperfective distinction. Recall that for the English simple present tense, a sentence like *John built a house* both implies that the goal had been reached and views the event as a

---

[115]This is done formally in (Lambalgen & Hamm, 2005, Chapter 11). In fact most of the types of coercion discussed in Chapter 2 are covered formally in this chapter.

[116]In such sentences as *She resembles her mother.*

[117]This concept will become a bit more clear when we discuss the computational machinery in the later sections. For now, it is fuzzy and intuitive. Think of it as possible substitutions to fit the reified fluents into a scenario. In this case only this possibility works, as is explained on (Lambalgen & Hamm, 2005, p. 174). The other, reverse substitution possibility fails because of the tricky *Releases* predicate.

[118]i.e. We want the complete event of 'build a house' to be the subject of a *Happens* predicate.

completed whole. In contrast, *John was building a house* implies no such completion of a goal, and views the event from an internal temporal perspective. The latter is easily formalisable. For example, the scenario contains the activity fluent *build*, which holds as long as the event is ongoing. So, we could represent[119] the past progressive sentence as

$$HoldsAt(build, t), \ t < now$$

which says that building activity was going on sometime in the past. Importantly, as the goal has not yet been reached at the time of the above statement, it is possible that extraneous events happen that stop the building process and the successful reaching of the goal.

For the perfective viewpoint, we need a way of utilising the scenario information in such a way that this is possible. Such a feat is accomplished by recourse to the hierarchical nature of events and levels of granularity discussed in the previous chapter. In this spirit what is needed is to fuse the various elements of the scenario into a single event. This is accomplished by defining a new event type. Let f be a fluent representing an activity, $start_f$ a starting event, and $finish_f$ the canonical terminating event, representing achievement of the goal. Define a new event type e by

$$Happens(start_f, s) \land Happens(finish_f, r) \land s < t \leq r \land HoldsAt(f, t) \ \rightarrow \ Happens(e, t).$$

Continuing with our concrete example, the perfective viewpoint of 'build a house' would then be

$$Happens(start_{build}, s) \land Happens(finish_{build}, r) \land \ s < t \leq r \land HoldsAt(build, t) \ \rightarrow \ Happens(e, t).$$

Should the *start* and *finish* events occur successfully, the temporal profile of the newly defined e will be set to the period equal between these two punctual events.

An activity such as 'push a cart' can be treated in the same way, but with $stop_f$ instead of $finish_f$, meaning that while there is no canonical goal to be reached, the activity is still gathered up into a perfective whole.

---

[119]Please note that this is not the proper way to do tense. That will become clearer in the upcoming sections. This example is for illustrative purposes only.

3.11   An informal sketch of the computational machinery

Now that the predicates, scenarios and different ways of constructing events has been explained, it is now time to sketch how they work. This is important, as while the predicates of the event calculus have intuitive plausibility, it is preferable to be able to fix their meaning more formally by relating them to each other via axioms. Secondly, we have been informally discussing comprehension in terms of discourse models. The use of axioms and a computational system allows one to explain how during comprehension of a sentence in an extended discourse that a discourse model is computed. The reader is directed to the appendix to see the full, formal definitions of the axioms, but many of them rely on the principle of inertia in slightly different forms. For our purposes the following informal formulation from (Hamm et al., 2006, p.15) is useful.

> if a fluent f holds *initially* or has been *initiated* by some event occur-
> ring at time t and no event *terminating* f has occurred between t and
> $t' > t$, then *f holds at t'*.

In even plainer English, this says that should a fluent start to hold at t, it will continue to hold until t' unless an *explicit* cause terminates the fluent before t'.[120] This can be demonstrated by a simple discourse.

(10)   It was hot. John took off his sweater.[121]

The first sentence serves as background, where 'hot' is true the whole time. It is a simple past sentence that denotes an event ('taking off a sweater') that is then inserted in the background state. This event structure interpretation is arrived at by the following reasoning. Our knowledge of the world generally says nothing about the effect of taking off a sweater on the air temperature. By the linguistic form ('it was hot'), we know that the state *hot* holds at some time t before *now*. By the predicates and axioms of the EC, it then follows that *hot* either holds initially (in our discourse model) or is initiated by some event. But there is no event specified in the discourse, and so *hot* holds initially. Importantly, since the discourse mentions no terminating event to *hot*, the principle of inertia tells us that it extends indefinitely. Therefore, the event denoted by the second sentence must be placed inside that of the first.

    Note that in the reasoning above it was implicitly assumed that as the discourse mentions only two eventualities, the event structure contains only these

---

[120]As can be seen by the above quote this principle of inertia also holds for those fluents that hold initially in the discourse model.

[121]The original example in (Hamm et al., 2006) is in French to demonstrate the discourse combination of the *imparfait* and *passe simple*. As the aspectual combination is the same for English in this sentence, there is no need to complicate matters any further.

two events. But this needn't be the case. For example, should the discourse have been continued with *Mary opened the window* the event structure would be different. A sentence would then be added to the scenario stating that the event of opening the window terminates the room temperature being hot.[122] This does not contradict the axioms, as they would only be violated if *hot* no longer held, but no terminating event occurred.

Thus it is the case that the axiom sketched above has the desired inertial effect only in 'minimal' models of the discourse. Again, resorting to an informal description in (Hamm et al., 2006, p. 15), 'minimal' refers to the requirements that

 i the model only contains those occurrences of events forced to be there by the discourse and the axioms

 ii the interpretation of the primitive predicates (*Initiates* etc.) is as small as is consistent with the discourse and the axioms

This means that in a discourse model no unforeseen events are allowed to occur and that the causal influences do not vary from what is expected. As hinted at in the analysis of the above mini-discourse, choosing to work with minimal models leads to non-monotonicity. In minimal models of the above discourse, the heat continues unless an additional event happens. Similarly, this sort of non-monotonicity is ideal for treating the English Progressive. Recall that a sentence like

(11) John was building a house.

may have a default implication to successful completion of the house, but it is not a strict implication. However, using the scenario from the previous section and the axioms, in minimal models the house will be built. But, should the sentence be

(12) John was building a house when a beam fell on his head.

an event is added to the discourse model that presumably terminates the building fluent and stops the canonical goal from being reached. Also, in our coercion example of a state used as an activity, we saw how the default discourse model is changed to yield a non-default interpretation.

3.11.1 Integrity Constraints

For the purposes of this interlude the final piece of the puzzle in describing how event structures are computed is to discuss integrity constraints. The relation

---

[122]This may seem like a common-sense bit of world knowledge, but even this notion is only a default. During the European heat wave of July, 2006, opening a window did not effect inertia one bit. Only industrial strength air conditioning would do the trick.

between a sentence and a discourse model can be viewed as follows. We want to view a sentence as a goal ('make S true') to be achieved by updating the discourse model. We can see how this works by examining the English present perfect. Take the following example

(13)   I have caught the flu.

where f is the fluent expressing *having the flu* and e is the event of infection. These two can be related in the language of the event calculus by means of the simple formula *Initiates*(e, f, t), which expresses the world knowledge that the event of infection is a cause of the flu. If we assume that this use of the perfect has present relevance,[123] then it is the case that I have the flu *now*. Given the idea expressed in the previous paragraph, let us see what it means to view (13) as a goal that says, 'Make "I have caught the flu" true in the given discourse model'.

Suppose we start with a pre-existent discourse model with facts about events and fluents. What is then needed is to construct a minimal adaptation of this model in which *HoldsAt*(f, now) is true. This cannot be done by merely adding this as a fact to the current discourse model,[124] but instead the sentence *HoldsAt*(f, now) triggers a kind of backward reasoning using the axioms of the event calculus, and scenario information (including basic world knowledge). In this case we can use the axiom of inertia discussed in the previous section.

We know that if the fluent f holds initially or has been initiated by an event at time t and between time t and t' no event occurs to terminate it, then it still holds at time t'.

Our conceptual information for this particular fluent says that it is initiated by an event e. In (13), no terminating event is mentioned, and so we can conclude by closed world reasoning that no such event occurred. Given this, and the axiom of inertia, there is only one fact to establish that I have the flu now, namely that the infection event e happened before *now*, i.e. *Happens*(e, t), $t < now$. This fact (and the logical consequence that occurs with *Initiates*(e, f, t)) is then added to the model. Nothing else need be added, but the result is that we end up with a model in which *HoldsAt*(f, now) is true. This abductive reasoning is what allows us to infer from (13) that there was an infection event in the past. This is the analogue of backward reasoning from goal to plan in robot planning discussed in the beginning of this interlude.

More generally in the event calculus, this sort of reasoning process is carried by a derivation procedure that in logic programming is called *resolution*. The derivation is started with a formula as a top query, which we want to make true in a discourse model. In the above case the top query is *HoldsAt*(f, now). To show that the purpose of a derivation is to test whether the query can be re-

---

[123]This is not always the case with the perfect. There is also the existential perfect as in, *I have caught the flu before. It's no fun at all.*

[124]As doing so may have unwanted consequences for the discourse model.

alised in a model, we write it as *?HoldsAt(f, now)*. The process then involves matching the query to the consequent of an axiom suitable to the query.[125] The initial query is then replaced by the antecedent of the axiom, which then becomes the new query to be resolved. As long as there are consequents of axioms to be matched, the process continues. It stops when there is a query that cannot be further resolved. In this case, it ends with *Happens(e, t)*, $t < now$. In the Event Calculus, this unresolvable query is interpreted as an instruction to update the model with material to make the initial query true. We will call this top query an *integrity constraint*, which is written as

$$?HoldsAt(f, now) \texttt{ succeeds}$$

Thus, we can now see from the above that the update nature of tenses is formally represented in the Event Calculus by means of integrity constraints.

Other types of Integrity Constraints

In the last section we saw an example of an integrity constraint that says that the query must succeed. This is not always the case and can be exemplified by the simple past tense when applied to non-statives. In these cases the English simple past is perfective and confines the event completely to the past. Compare

(14)   a. #John ran and he still is.
           b. John was running and he still is.

Taking the sentence *John ran*, the first attempt at formalisation would be to construct an event e via the coding methods and hierarchical planning (to capture the perfective viewpoint of the simple past) and have the integrity constraint be

$$?Happens(e, t), t < now \texttt{ succeeds}^{126}$$

This places a *run* event in the past, as desired, but does not rule out the event continuing into the present in our model, in contradiction to the linguistic fact seen in (14). In minimal models the set {t | *Happens*(e, t)} consists of a finite set

---

[125]In this case the axiom of inertia for instantaneous change will do. Formally, it looks like *Happens(e, t)* ∧ *Initiates(e, f, t)* ∧ $t < t'$ ∧ ¬*Clipped(t, f, t')* → *HoldsAt(f, t')*. It is the *HoldsAt* in the consequent of this formula that we match our query to.

[126]This is inaccurate for two reasons. The first will be discussed below. The second is that instead of anchoring the time to an instant t, it is actually anchored to a Reichenbachian reference time R. For details, see (Lambalgen & Hamm, 2005, Chapter 8).

of intervals. Each interval is an e token.[127] We want to make sure that there is one token that lies completely in the past. This is done by a negative integrity constraint.

$$?Happens(e, now), \texttt{fails}$$

which says that any update resulting in *?Happens*(e, now) is not allowed. The simple past then has two integrity constraints, one positive and one negative needed to capture its update properties. They can, however, be combined in to one constraint[128]

$$?Happens(e, t) \; t < now, \texttt{succeeds}, \; Happens(e, now) \; \texttt{fails}$$

---

[127]This is a bit complicated. For the full explanation the reader is invited to see (Lambalgen & Hamm, 2005, p. 41-43). But think of events in terms of fluents. Interpret fluents as the sets of intervals $(a,b]$, where a is the instant where an initiating event occurs and b the instant where the next terminating event occurs. We can then define event-types in terms of the set of intervals (for a given fluent) where, speaking intuitively, this fluent is intimated or terminated. A given interval in this set is an event-token (i.e. an actual event occurrence.)

[128]And give us something that looks like the proper representation for the simple past.

Chapter 4

Perception Verbs

Who are you going to believe, me or your own eyes? – Chico Marx, pretending to be Groucho

4.1 Introduction

This chapter is an examination of the logical, syntactic, and semantic properties of English verbs of perception, such as *see*, *hear*, etc. For the present, two classes can be differentiated – a more stative class and an activity class. In English, verbs for visual and aural perception have the pairs *see/watch* and *hear/listen* that differ in their aspectual properties. Both types shall be explored, but emphasis is to be on the former type. Aside from the aspectual difference between the *see* and *watch* type, verbs of perception – primarily *see* and *hear*, have a number of other senses that go beyond direct perception. The examples below illustrate this.

(1)   a. I saw Mary leaving the room.
       b. I saw that Mary had left the room (as her coat was gone).
       c. I see now that global warming is not an anti-capitalist conspiracy created by loony environmentalists.
       d. I hear what you're saying, but I don't agree.

Sentence (1-a) is a direct perception sentence, that describes my perceptual experience of the event of Mary's exit. (1-b), while perhaps involving direction perception (seeing a closet without a coat) is nevertheless indirect perception as the exit event is not seen, but inferred from the perception. Sentences (1-c) and (d), show the use of perceptions verbs to express understanding or a realisation. Only the direct perception type will be investigated, here.

Section 2 explores the logical and psychological properties of direct perception reports – examining the issue of whether if something is seen is enough to then infer that the event actually happened. That is, are direct perception reports

information preserving? Section 3 examines the syntax of direct perception reports as well as the non-logical semantic properties, especially the temporal and aspectual interaction between the matrix verb and complement. A recent minimalist account (with a Kratzerian event semantics link) of (Felser, 1999) is then examined. This account, however, in the author's view fails in two main areas. The first is that the coordination of tense between matrix and complement (that, e.g. the seeing of the matrix and the event happening in the complement must be simultaneous) is treated syntactically. Doubts are cast upon this by an examination of other complement taking verbs where the simultaneity property does not hold, leading toward a semantic treatment of tense coordination. Secondly, the aktionsart of *see* is actually a rather complicated matter, and needs a more detailed examination than given in Felser's account (or any other that I am aware of).

This is the subject of the fourth section, where the aktionsart of both *see* and *watch* are worked out in great detail, as well as the rather complex interactions that occur when varying tense and grammatical aspect are paid close attention to. With the tense, aktionsart and aspectual phenomena accounted for, Section 5 is a formalisation of both *see* and *watch* that captures the varying tense and aspect, as well as the logical properties taken as a strong default, rather than strictly information preserving.

Finally, there is a brief look at the use of *ing* complements, and the rather complex way they receive tense, depending on verb type and context. This both reinforces the need for a semantic treatment of tense coordination in direct perception reports, and leads into the study of *ing*-nominalisation that is the subject of the following Chapter 6.

4.2 Logical and psychological properties of perception verbs

Before examining the tense and aspectual properties of English perception verbs, an excursion into the well-known logical properties is first necessary. While both the psychological and logical properties of these verbs are not the main focus of this chapter, insights gained from these will inform the necessary modelling in the event calculus. We shall first briefly look at the seemingly natural logical characterisation of perception and its presumed implications. Next, some objections from (Lambalgen, 2002) and (Jackendoff, 1992), showing that a look at the psychology of vision sheds light upon its supposed logical characteristics. Finally, observations from (Gee, 1977) are used to show that both psychological aspects of vision as well as cultural knowledge influences the use of perception verbs. Indeed, it is the ordinary language use of verbs such as *see* that is at issue here, not a (most likely hopeless) account of direct perception of the world that is then harnessed by various perception verbs. The above may seem a bit far afield

for one whose intention is to study tense and aspect interactions in perception reports, but it is needed as a starting point of the formalisation of *see* and related perception verbs.

4.2.1 Logical properties

The three listed principles are the hallmark logical properties of perception reports.[1]

1. Veridicality: For simple[2] sentences, if a sees ϕ, then ϕ

2. Substitution: If a sees $\phi(t_1)$ and $t_1 = t_2$, then a sees $\phi(t_2)$

3. Exportability: If a sees some x such that $\phi(x)$, then three is an x such that a sees $\phi(x)$

For now, we shall focus only on the first and then come back to some interesting thought experiments by Gee that questions that viability of substitution. The first principle – veridicality – says that the complements of perception verbs in a direct perception context (i.e. bare infinitives and the corresponding *-ing* version) are factive. If I saw Mary cross the street, then indeed Mary crossed the street. Of course the converse also holds – that is, if Mary didn't cross the street after all, then I couldn't possibly have seen it.

This view takes perceptual reports to be information preserving. What is seen is the "objective" world[3] and so a semantics of a verb like *see* should be formulated in terms of this objective world.[4] Therefore, the notion of veridicality must go through. If I perceive Mary walk across the street, and perception is that of the objective world then the utterance, "I saw Mary cross the street" must indeed imply that Mary did such an act. Contrast the use of *see*, with a non information preserving verb such as *believe*. I can believe that the earth is flat, with no implication that it actually is. However, on the veridical view of *see*, I could not have seen Mary cross the street if she actually didn't.

---

[1]The below formulation is from (Higginbotham, 1983), but similar can be found in much of the semantic literature.

[2]'Simple' is meant to exclude irrealis contexts.

[3]Or actually, bits of it at any one time.

[4]Substition allows for the replacement of different names or descriptions of the same object to be substituted *salva veritae*. For example if Mary is also my sister, then if I saw Mary walking I also saw my sister walking. Were this a case of indirect perception, e.g. I saw *that* Mary was walking, but didn't know she was my sister the substitution would not go through. Exportability means that if I see Mary where Mary is walking, then there is such a Mary such that I see her walking.

There is certainly reason to take such a view. In ordinary life, we are often more convinced evidentially by a perception statement, providing the witness is reliable. It is usually only in circumstances such as bad light, fog, a house of illusions and the like that one is liable to question the content of the report. In criminal law an eyewitness report of what happened is considered to be rather strong evidence, often the strongest type.

However, there is a tension between "truth" of the complement and what is seen. An attorney cross-examining an eyewitness report can attack its veracity on two counts. The first would be that the witness is a liar and cannot be trusted. There is no issue of veridicality here, and (if a lie) the perception report can be thrown out as unsafe evidence. The second method of attack is more interesting for our purposes.

An attorney can cast doubts upon the content of the perception itself. If it was a dark and foggy night, was the witness *certain* that it was Mr. X that she saw robbing the house? Did she see anyone at all, or was it a trick of the light? This case is more interesting than the first because the witness is perfectly justified in her perception report – she merely told the court what she saw. The defense attorney would couch his language[5] in such a way to deny that the witness saw what she said she saw. *Our esteemed witness* claims *she saw Mr. X robbing the house, but this is plainly impossible...*

Assuming the defense attorney is correct, what can we say happened? The lawyer would say (using the principle of veridicality and *modus tollens*) that the witness did not and could not see Mr. X rob the house, because Mr. X did not rob the house. But yet the witness is honest and when presented with Mr. X in a line up picks him out as the culprit she saw. It turns out that the house was never robbed to begin with and tricks of light, fog, and an over-active imagination led the witness to her perceptual experience. Indeed, after argument and reflection, the witness may indeed agree that she did not see what she thought she did. Yet, at the time of the supposed crime she was perfectly justified in saying, "I see a man robbing the house", because this is exactly what she perceives at the time.

Moving away from courtroom drama and contrived stories, ordinary language has a simple way of exhibiting the tension between perceptual reports as information preserving and as a mere report of what the individual perceives.[6] There is the expression, "I was just seeing things" to describe the experience of "seeing" something that isn't there. In this expression, not only is veridicality not assumed, it is contradicted. The expression contains within it the denial that anything was seen, but does not deny the *seeing* itself.

The above stories regarding witnesses and the supposed veridicality of what they see, focusses on the power of reports of perception, not the act of perception

---

[5]All decent attorneys make the interpretation procedure as difficult as possible.

[6]This is already assuming that one does not directly perceive the world, an inarguable point for several hundred years already.

itself. In ordinary life, perception itself is of necessity 'veridical'. That is, if I am walking down the street and see a tree, I do not think about what I am seeing and do not for one second doubt the existence of the tree. Certainly, I would not purposefully walk into it, just to find out if my perception is non-veridical. Perception is just my way of being in and being able to interact with the world,[7] and generally not something doubted, or even thought about. This certainly supports a strong veridicality for perception.

That said, the examples above show that in ordinary life, we know that our eyes can sometimes deceive us and even acknowledge this with such expressions as the cliche in this sentence. Moreover, we learn through experience the (ordinary) circumstances wherein our perception does fail, such as bad light, fog, bad vision, or simple optical illusions. It seems that we move though the world as if our perceptions were indeed veridical[8] we are acutely aware that it is not always the case, and this doubt does show up in perception reports.

We are left, then, in a bit of a quandary. On the one hand, the philosophers and semanticists do not have a monopoly on the notion of truth preservation and veridicality – there is a strong tendency in many cases to assume such principles in ordinary life as well. However, the brief examples above show that ordinary language recognises that not all seeing is veridical. It is perfectly reasonable to say that I saw something that wasn't there. Should a drunk come up to me and exclaim over the wonder of the dancing pink elephants he saw, I would doubt the existence of such fanciful elephants wandering the streets of Amsterdam, but I would be inclined to believe that he saw such a thing. One way to dissolve this tension is to assume that the relation between perception and the world is mediated by psychology, as well as cultural practice and concepts. With these interferences, it is quite plausible to imagine that people actually *see*[9] things that aren't there.

4.2.2 Psychology and perception

One way to account for both the compelling veridicality of perception reports, but still holding this to be a default is a move from direct perception of the world to mediated perception via psychological modelling. This approach has been taken in various forms in (Jackendoff, 1992), (Miller & Johnson-Laird, 1976) and (Lambalgen, 2002). For example, the 'visual experience' one has may change given additional information, thus effecting what one sees. For example, while

---

[7]I am concentrating on visual perception here, because at least for the sighted, this must be the most 'compelling' sense.

[8]In fact, it would be rather difficult to live in the world if we acted as if the world were a funhouse or a mass hallucination. I believe that it was Hume who commenting on Berkeley's arguments for scepticism noted something to the effect that while they were rationally unimpeachable, they were not so powerful that one would be compelled to walk of a cliff.

[9]Not merely *think* they see.

walking in the forest I can see a snake on the ground, but while carefully moving closer, it is only an old rubber hose. There is no change in the object, of course, but my visual information is refined, changing what I see. While actual and more sophisticated psychological accounts can be found in the above references, the intuitive little story is to show that accounts of *see* have used a causal process to mediate between the perception report and the world. When this mediation happens to match what is in the world, verdicality follows.[10]

While in no way doubting that some variation of this causal story is what actually goes on in perception, I have my doubts as to whether it captures the meaning and usage of it.[11] Instead, I have taken some inspiration from (Wittgenstein, 1953, p.193 - 214) for a way out.[12] His remarks on *see* seem to be a reaction to a sort of Russellian empiricism about vision. This view can be simplistically characterised as follows: when I look at a box, I receive sense data that are then put together in such a way that I interpret the image in my mind as a box.

Wittgenstein is sceptical both of a causal explanation of perception to account for its meaning,[13] as well as the role interpretation plays in perception. Unarguably, interpretation plays a role when critiquing abstract art, seeing figures in clouds, or a children's game where the child sees the box as a castle. But in what we may call 'ordinary' perception this is not the case, rather the perception is not interpreted, but just something that forces itself upon us. Wittgenstein gives two examples, one with an optical illusion that changes aspects[14] and that of an ordinary object. On page 194-195, he illustrates a situation where someone is shown the figure for the first time and is unaware of the trick.

> I may, then, have seen the duck-rabbit simply as a picture-rabbit
> from the first. That is to say, if asked "What's that?" or "What do
> you see here?" I should have replied: "A picture-rabbit"...I should not

---

[10]Thankfully, this often does seem to be the case.

[11]I should say here that I am not rejecting all use of psychology and cognition in relation to semantics. The introduction, and first two main chapters make much use of psychology and general cognitive principles in regard to general event structure, the interpretation of tense and aspect and the like. Moreover, I think that a certain folk psychology will certainly have an impact on the usage of various lexical items, including verbs of perception. But, I don't believe that a causal account alone captures the meaning of a lexical item. Not that I am accusing any of the above authors of going that far.

[12]The word *inspiration* should be duly noted. I am not going to (or am able to) give an exegesis of these confounding, but beautiful remarks. I am still trying to grasp them fully and would not want to claim Wittgenstein holds something that just happens to be a product of my misunderstanding.

[13]*Its* causes *are of interest to psychologists*, he notes on p.193 in regard to noticing a new aspect when looking at a face.

[14]The famous gestalt duck-rabbit picture that can appear as a picture of a rabbit or picture of a duck, but not both at the same time. Once one is familiar with the trick and has a little practise, they can then witch between the two images at will. This latter switching between aspects consciously would be a case of a more 'interpretive seeing.

> have answered the question "What do you see here?" by saying: "Now
> I am seeing it as a picture-rabbit". I should simply have described
> my perception: just as if I had said "I see a red circle over there"
>
>   Nevertheless someone else could have said of me: "He is seeing the
> figure as a picture-rabbit."

The last line is also quite important as it exhibits the differences that can arise between first and third person perception reports.[15] It also shows that when someone is aware of the trick, they can bring a notion of interpretation or awareness in the perception report. This is, of course, not only possible in third person reports. Once I am aware of the aspect-switching trick, I can have hours of fun saying, "Now I see it as a rabbit, ooh now a duck, etc." If I do not know the trick and merely see one of the animals on a consistent basis, this makes no sense.

Sometimes, this is not the case. In continuation of the quotation above, Wittgenstein gives an example with a example of completely ordinary object perception.

>   It would have made as little sense for me to say "No I am seeing it
> as..." as to say at the sight of a knife and fork "Now I am seeing this
> as a knife and fork". This expression would not be understood.–Any
> more than: "Now it's a fork" or "It can be a fork too".

In this case, there is no room for thought or interpretation in the perception of the cutlery. One simply sees it, and doesn't interpret or take it to be something.[16] Here, the third person observer would also not be able to say, "He is seeing it as a fork." There is just a fork and nothing else to be spoken of. However, we can use the split between first and third person reports to get a grip on a default notion of veridicality with a simple illustration.

I see a monster

A child is sleeping in his room at night, wakes up and looks around. While the light is still dim and his mind still leaden with sleep, his eyes turn to a coat on a chair. He becomes terribly frightened, and screams out.

> *A monster. I see a monster in my room*

---

[15](Lambalgen, 2002) do note the difference between more epistemically neutral third person perception reports and those in the first person.

[16]It is the context, not the content of perception that seems the most important. The responses that make little sense at a dinner table would be perfectly acceptable at an exhibit of Marcel Duchamp's long lost work of dadaist table-setting.

His parents rush in, turn on the light and explain to the child that the 'monster' was merely a coat hanging on the back of a chair. Assume that the child realises this, and calms down. He could then say any of the following.

(2)   a. I thought I saw a monster.
         b. I saw the coat and thought it was a monster.
         c. I saw a monster but it was really my coat.

and so on. His parents can also utter the equivalents of the above, but in the third person. I would say both past tense versions (first and third person) are both a report *about* direct perception, that includes the result of a conscious interpretative process.

The parents (suppose they were watching the action from a webcam), can describe what happens in the present tense as well, but it is still an interpretative process that is about their child's perception, and not a report of what they are perceiving directly. Having greater knowledge and observing their son's behaviour (looking at the coat and cowering in fright), they could say any of the following:

(3)   a. He sees a monster.
         b. He sees a monster, but it is really a coat.
         c. He sees a coat and thinks it's a monster.
         d. He sees the coat as a monster.

Though in real present tense, these statements are again the result of an interpretive process and not a perception report of what they are seeing. I would say that the only true direct perception reports are those of first person present tense. All others will be an report of someone else's experience (with or without an interpretative leap) or a report of one's own perception, but one that happened at an earlier time.

With these less direct perception reports it is possible to both say what was being seen (in the sense of the experience) and the facts of the matter, if they should differ. For both the third person or first person past, either one of these cases will be that of a report about a direct perception, that is mediated by an interpretive process. Even in the present tense, when the parent says, 'He sees a monster', this is because of their interpretation of their son's behaviour. When the son says, 'I see a monster' this is purely his experience.

Importantly, I would say that though there is no monster, both (2-a) and (3-a) are true statements, though with slightly different content. The child has an experience of a monster, he does briefly *see* one even though it doesn't exist. The parents are certain there is no monster, but they know the pitfalls of perception and can adjust their perspective accordingly. They may have even had a similar experience as a child. More generally, I think this sort of folk psychology can colour our use of perception verbs and their logical properties without relying on a more sophisticated causal explanation. Often, when we make a perceptual

mistake, we are able to reconcile this with something in the world that was the object of perception, as such a thing has happened to us many times. In the most extreme case, we have all hallucinated[17] and know that we can see things that have no basis in reality whatsoever and this is also reflected in our language. So, while *see* has a very strong notion of veridicality, there are cases that we are aware of where it can be denied, and this does need to be taken in account in the semantics. The question is to how this can be accomplished.

4.2.3 Preference Rules

Jackendoff, assumes that visual perception consists of a sort of psychological modelling process, and that while veridicality may not be an absolute implication of any perception statement, there is a strong tendency to assume it is so. It is usually unordinary circumstances or reports of strange objects or events that cause one to wonder at a perception statement's veridicality. Indeed, the truth of the complement of a perception report should be assumed as a default.[18]

To account for both intuitions, (Jackendoff, 1992) employs preferences rules to account for what appear to be two different notions of *see*. Take the following examples from (Jackendoff, 1992)

(4) a. Bill saw Harry.
     b. Bill saw a vision of dancing devils.
     c. Bill saw the tree, but he didn't notice it at the time.[19]
     d. #Bill saw a vision of dancing devils, but he didn't notice it at the time.

(4-a) is a prototypical perception report where veridicality is a rather strong default, whose veridicality is difficult, if not impossible to cancel.[20] Though we know (or at least hope) there are no dancing devils, (4-b) is a perfectly reasonable report as well. Jackendoff modifies matters somewhat by using 'vision' as a mediator, but (b) and (d) without 'a vision of' are still felicitous and infelicitous respectively.

Sentences (4-c) and (d) shows what happens when the default is cancelled. But before explaining the differences between (c) and (d), as well as some subtleties not noticed, the predictions in (4) are explained in (Jackendoff, 1992) by preferences conditions for *x sees y*.[21]

---

[17]Sometimes from illness, sleep deprivation or other reasons.

[18]We have already seen that the strength of this default can depend on both the tense of the report and whether the report is in the $1^{st\cdot}$, 2_{nd} or 3rd person.

[19]I find this rather odd. If Bill saw the tree he wouldn't have run into it, which makes me worry about the sentence's felicity. Perhaps, he meant to say that Bill saw *something*, but didn't realise at the time that it was a tree.

[20]This assumes the statement is in a honest, omniscient third person. A narrative that combines a first person point of view told with the third person would be more doubtful. Should someone like Alain Robbe-Grillet have written that line, I would assume the opposite.

[21]Preference rule systems in general are Jackendoff's way of capturing the default nature

1. x's gaze makes contact with y.

2. x has a visual experience of y.

A felicitous use of *x sees y* need only have one of the conditions satisfied. If both are, as presumably is the case in (4-a), it is prototypical, veridical seeing. There is presumably nothing to gaze at in (b), but Bill's experience is enough to justify its felicity. (4-c) only satisfies the first condition as the second condition is explicitly refuted, but is still passable. Only when neither condition is satisfied is the use of *see* infelicitous – as happens with (4-d).[22]

However, there is a major problem here. We have all had the experience of looking at something without seeing it. For some reason, simple items like keys seem to disappear and then reappear in the same place after having looked there for the fifth time. Yet on these preference rules, I would have seen the keys just by looking at them. This is similar to the example (4-c), which hides a good deal of subtlety. Modifications need be made to Jackendoff's preference rules to flesh out the conditions necessary for perception reports.

4.2.4   Modifications

Complexities of object seeing

Imagine (4-c) uttered by Bill himself, in the present tense and past tense respectively.

(5)   a. ??I see the tree, but I don't notice it.
         b. I saw the tree, but didn't notice it at the time.

The situation could be one where our evolutionary defects are exposed – being absentmindedly lost in thought and walking into a tree while looking straight at it.[23] Certainly, if Bill was in a position (cognitively and conceptually) to utter, "I see a tree in front of me", it is doubtful whether Bill would have walked into it.

---

of concepts and try to capture such notions as Wittgenstein's *family resemblance* notion, and avoid needless proliferation of senses.

[22]While the notion of 'visual experience' seems nicely intuitive, thinking about it deeper leaves me confused. (Wittgenstein, 1953) has a cryptic (for me) passage differentiating between seeing a rabbit sitting and having a rabbit suddenly run past. In the first case, 'I see a rabbit' is a report of perception, in the second it is an expression of visual experience, something forced from us like a cry is when we are in pain. I prefer his caveat, "Do not try to analyse your own inner experience." I attempted the opposite, which I would not recommend. I walked down the street and tried to examine my visual field as something separate from myself, picking out various bits of it that I would recognise. Within twenty seconds, I ran into a pole.

[23]Sadly, this has happened to me on an all too frequent basis.

Bill's gaze did make contact with the tree, but he had no visual experience of it. Using Jackendoff's criteria, the gaze condition alone should make *I see a tree* felicitous. Yet, in the scenario given above, this is plainly absurd. In the retrospective perspective of the (5-b) version, this becomes more plausible. Bill could mean that, looking back, there was some dark area in his field of vision that was the tree, although he wasn't paying enough attention to realise he *was* seeing a tree. (5-b) is still a bit odd, but passable, taken in this sort of clarification sense. Of course, one can also look straight at an object and see nothing at all,[24] which might cause one to question the felicity of even (5-b) in this case.[25]

Combine this with the previous section's casting of doubts on the veridicality of first person seeing, and it is plain to see that the preference rules do not quite work. Rather, with present tense, first person direct perception felicity is dependent only upon the visual experience condition. I can see a ghost or devil only because I have a particular visual experience and my gaze is not directed at these non-existent objects. As was just shown, it is quite possible to have one's gaze directed at an object and have no visual experience whatsoever – something that is not a sufficient condition for first person, present tense seeing. Moreover, it must be the case that if I see X, then the visual experience of X is recognised as such, and this is where conceptual knowledge plays a role.

With a visual image of a tree in mind, this might seem a rather trivial point. This, I believe, is also what helps (5-b) make sense. One knows how to see a tree, and can introspectively realise that this is what he was looking at, though he didn't realise it at the time. But suppose I am learning elementary biology and a tutor is guiding me through the basics of microscope work. Looking at a slide that supposedly has fine examples of a particular virus. Being a novice, when I look through the microscope I see very little I can define, and certainly nothing I recognise as a virus.

At this point, if I tell my tutor that I see the virus, I am not only lying, but doing harm to my biological education. Telling my tutor that I see the virus would lead her to believe that not only do I now grasp microscope viewing basics but also *how* to see the particular virus. Here is where matters become interesting. Suppose I am a biological failure, never learn how to use a microscope, or how to see a virus through one. Even retrospectively, as I still don't know what it is to see a virus, it is rather odd for me to say

(6) ??I saw the virus, but didn't recognise it at the time.

As a more honest student, I tell my tutor I am having difficulty with the microscope process, and she dutifully guides me through the process. Even though I

---

[24]There is the common experience of looking for one's keys in the same place many times and finding nothing. Suddenly, they appear as if by magic.

[25]It may be a case of idiolect differences, but I would hesitate to use the verb *see* in either (5-b) or (4-c), and instead employ the less conceptually loaded, *look at.*

still see nothing but random shapes and colours, she can use the 2_{nd} person form to tell me, "What you see in the centre of your visual field surrounded by the black dots is the virus."[26] I still may not recognise it and still not see it, but after reading a nice picture book, learn how to see it. Now, (6) is perfectly felicitous. I leave it to someone with a better intuition than myself to decide whether the two cases sketched above affect the felicity of the third person version.

This example may seem fanciful, but is actually rather common. Being a fan of detective novels, I have read many times the detective being frustrated when confronted with forensics. Often, when dealing with the coroner or scientist about a piece of evidence, they will sadly wish that they could be able to see what the forensic scientist does. While I can't say I understand this phenomenon completely, I don't think it is an example of interpretation in perception, but a skill on the part of the forensic scientists that changes how he sees (in certain contexts) once he has become involved in a specific practice.

Finally, as with the keys example, it is quite common to not see or hear objects or events that are quite ordinary. For example, I live very near a church where loud bells ring every fifteen minutes. Thankfully, after a week, I rarely heard them, though doubtless they are still ringing with appalling regularity.

Conceptual affects of activity and event perception

With direct perception of events and actions, the perceptual content or how it is too see an object, may not come into play at all. Similarly to objects, the same physical activity can be described in a number of different ways. The deciding factor in what an activity actually is may be only the intention of actor, conventional rules, or background knowledge the speaker does not have to hand. Slightly altering an example from (Gee, 1977)

(7)   I saw Tom steal your car, but I didn't realise it at the time.

At issue is not that of substitution[27] – the speaker saw Tom doing something – but rather the activity seen. I see Tom use keys to unlock my friend's car and drive away. At this point and my background knowledge,[28] I see Tom borrowing the car. It is only later that I learn Tom and my friend have had a row that included Tom taking the car keys and stealing the car. My friend could say to me, "You didn't see Tom borrowing my car, you saw him stealing it." Our omniscient third person narrator could say, "He saw Tom stealing the car, but at the time thought it was only borrowing."[29] Only once I know that Tom has

---

[26]This is a thought experiment and not meant to be an accurate description of viewing a virus in a microscope – something I know very little about.

[27]If Tom is my brother, substitution would imply that the intensional transparency of *see* demands that I also saw my brother steal the car.

[28]It is a mutual friend's car.

[29]Present tense works perfectly here as well.

stolen, and not borrowed the car, can I utter sentence (7).

The same veridicality issues with respect to person and tense occur, but for different reasons. It is not the refinement of perception or learning how to see a particular activity that is at stake here,[30] but an update of knowledge that changes what the activity seen actually was. Stealing and borrowing can (as a physical activity) take many different forms. Certainly, if one had a gun, or broke a window intentionally, one would more likely infer the former, but sometimes only context can differentiate between the two. The past tense, first person version with the correction is only felicitous once the actual act is known. This parallels the way the object version (*I saw the virus, but...*) is only possible once I know how to see a virus.

Notice that interpretation is certainly an issue here, but in a strange way. In the above example, what is seen, spatio-temporally, does not change, but our definition of what the event is, does. But, as shown in Chapter 3, intention and social rules are part of what defines events just as much as spatio-temporal changes, and that is what is at stake in this example – not the perception, but the proper reporting of it.

Substitution issues

The rather common example of substitution when applied to activity seeing is as follows. Assuming Carol is the mayor, then if I see Carol jog in the park, I also see the mayor jog in the park. Of course, if I did not know this identity at the time, I may not know or utter this equivalence. On retrospect, and now knowing that Carol is the mayor, I would assent to the fact that I saw the mayor jog in the park.

Most ordinary examples of both object and activity seeing would follow along these lines. But, as hinted at earlier, there is something rather special about *seeing* that allows exceptions to this truth preservation principle. (Gee, 1977) uses Wittgenstein's drawing of the duck-rabbit and notes that even if the observer knows that duck=rabbit, substitution does not work. Suppose that while staring at the picture for a bit too long, I see the duck wink. No matter the veridicality, it is certainly reasonable for me to exclaim, "I saw the duck wink". However, even if I am familiar with the illusion, and have seen the rabbit before (and have even see the rabbit wink, before), it would certainly false for me to utter, "I saw the rabbit wink," as I saw no such thing. Recall the conceptual nature of seeing has a certain immediacy and (for lack of a better word) intimacy, especially in the ordinary case, where the visual experience is something that happens to us, is impressed upon us. From (Gee, 1977) (p.477)

Semantically, I believe NI-constructions[31] have a particularly close

---

[30]Though this could also be the case – there are plenty of hopefully apocryphal stories that end with a small child asking, "Why is daddy hurting mommy?"

[31]'NI' means 'naked infinitive' and refers to what are called 'bare infinitives' in the next

relationship between the VP in the complement and the higher perception verb (an almost "direct object"-like relationship. In the way in which "John felt Mary" means that what John felt was Mary, "John felt Mary hit him" means that what John felt was the hitting of Mary on him.

Whether or not it is even correct to say that duck is the same as the rabbit in the picture, it is certainly incorrect to say that the activity of a duck winking is the same as that of a rabbit winking. In this light, it is rather that even explicit knowledge of the identity of the two picture animals is not enough to save substitution. Certainly, this relies on the fact that the duck and rabbit cannot be seen simultaneously. Unlike the example of Carol and the mayor, which is merely an issue of world knowledge, seeing a duck and seeing a rabbit are two completely different experiences. It would certainly follow that seeing activities or events involving them would also be completely different experiences.

4.2.5   Conclusion

Our initial notions of perception reports is that a perception verb is truth preserving. After all, we are seeing the world, and so if we honestly report seeing something, it must be the case in the world. A few examples and a little psychology of the perceptual system are enough to show that this is not quite the case. There is a certain rhetorical power to perception statements that makes one think that in general, we do take them to be truth preserving. Indeed, certain language styles such as third person narration indeed lend themselves to perception statements being truth preserving.

But when examining first person statements, especially in the real present tense, many factors can go awry – perceptually, conceptually and contextually. These caveats apply where the complement of the perception verb is either an object or event. Truth preservation, instead, must be taken as a default of a perception statement – the strength of the default depending often on the grammatical number and tense of the report.

We saw one attempt by Jackendoff to give default conditions for *x sees y* via a pair of preference rules, only one of which needs to hold for such an utterance to be felicitous. This is certainly on the right track, but breaks down, especially in the case of first person perception reports. Ultimately, we need to look for conditions that extend beyond an individual's psychology,[32] and add conditions based on a more conventional, conceptual readiness that emerges in a individual as part of the ordinary social and cultural learning process.

---

section. They are sentences of the type, "I saw the duck wink" as opposed to the participial, "I saw the duck winking."

[32] A theory of an individual's psychology is necessary, but not sufficient for a full account of perception reports.

I have no equivalent of Jackendoff's preference rules that could handle many of the varieties of perception reports, but it is clear that there needs to be some sort of perceptual readiness.[33] Secondly, we must assume veridicality as a strong default, that if mistaken[34] has something in the real world that it can be matched to. The limiting case would be a hallucination, where there is nothing in the real world for the hallucination could correspond to.

4.3 Perception reports of eventualities

In the previous section we examined various logical and conceptual properties of perception reports – mostly of object perception. It is now time for a slight change of direction, and examining the tense and aspect interaction involved in event perception. We will first look at basic syntactic, aspectual and argument structure properties of direct perception sentences. This is followed by an examination of the status of the activity complement, with special attention paid to recent work in a minimalist framework by (Felser, 1999). The following brings a deeper examination into the interaction of aspect and aktionsart of both *see* and *watch*, leading to a formalisation of these interactions in the Event Calculus.

But to start it is best to constrain the class of perception reports we are to examine. Direct perception of activity, such as

(8)   a. I see John walking across the street.
       b. I hear a bell ringing.
       c. I'm watching John walk across the street.

brings up many interesting tense, aspect and aktionsart interactions on top of the general logical and conceptual difficulties inherent in perception reports. Examining the relevant interactions in perception verbs sheds an interesting light on the conceptualisation of events. Moreover, there is Wittgenstein's 'untangling' of *see* into a variety of related senses, coupled with the intimate relation between the 'act'[35]of seeing and what is seen. Perhaps, then, it is better to say that the perception report itself takes a number of different event-type senses depending on how the activity is characterised.

---

[33]With the caveats that sometimes our senses may work perfectly but still we don't see or hear something. It also needs to allow for cases where an expert in a field is able to see things that an amateur can only dream of.

[34]For an event, it could be the stealing/borrowing example, or even mistaken spatio-temporal perception such as seeing a monster tapping a the window that is actually a tree branch blowing against the window in a storm.

[35]Wittgenstein rightly characterises *seeing* as a state, but in the sense of a mental state, not in the event-type sense of the word. Nevertheless, it will be seen that with verbs such as *see* or *hear*, the nature of the activity (in its aspect and event type) does colour (phenomenologically) the perception itself.

We will begin by looking at the 'syntax'[36] of perception reports and, as it becomes increasingly necessary, bring in certain semantic notions. This will be followed by an examination of the necessary aspectual interactions between matrix verb and complement that make verbs such as *see* and *hear* rather special, and a challenge to formalise.

4.3.1   Syntax and Argument Structure

We first need to define what is meant by direct perception of an activity. The two basic types are as follows.

(9)   a. I saw Mary crossing the street. (participial)
         b. I saw Mary cross the street. (bare infinitive)[37]

The first takes an *-ing*, participial complement and the second a bare infinitive (*crossing the street* vs. *cross the street*). Both complements have no tense marking themselves, something that turns out to be very important for their understanding, both syntactically and semantically. With the related, indirect perception sentences, the complement carries tense and will sometimes have a marker such as *why* or *that* emphasising that the eventuality is independent of what is seen.

   Indirect Perception Sentences:

(10)   a. I see (that) Mary crossed the street.
          b. Jane saw why Peter had voted Green.

Intuitively, what is seen in the second case is not the eventuality, but in the case of (10) (a), some evidence for the result. For example, if I see Mary on one side of the street and then the other a minute later, (10) (a) is quite a reasonable inference. However, neither (9-a) or (b) would be a felicitous account of what I saw. Both (10) (a) and (b) are rightly called indirect as no direct perception need be involved in the utterance. It is a use that plays upon the importance of understanding or realising in direct perception, but extended to interpretation and inferences. The sceptical reader should attempt to replace *see* with *understand* in both sets of examples to see the differences for themselves. This chapter is concerned with sentences such as those are seen in (9).

---

[36]*Syntax* is in quotes, as syntax will be of a mostly informal nature in this chapter. The reader will be referred to works that are more concerned with and focussed on a formal syntactic account of these structures.

[37]I will follow (Felser, 1999) and use the terms PPVC and IPVC to be shorthand for participial perception verb complement and infinitival perception verb complement.

Catenative Verbs

The unique character of perception verbs can be seen, when we compare them to other complement taking verbs. At first glance, there seems to be a syntax/argument structure mismatch. In (Huddleston, 2002a), *See* or *hear* are verbs that come under the heading of *catenative* verbs, i.e. verbs that take complements as well as ordinary nouns as objects. Basically, there are two common types of catenative verbs, the *catch* and *regret* type, in the sense of how the syntactic arguments align with the semantic ones.[38]

(11)   a. I caught <u>Kim</u> mistreating my cat.
        b. I resented <u>Kim</u> mistreating my cat.
        c. I saw <u>Kim</u> mistreating my cat.

In sentence (11-a), Kim is the argument of *catch* – Kim is caught in the *act*, one could just as easily (though a bit awkwardly) say, *I caught Kim while she was mistreating my cat.*

    With sentence (11-b), Kim is not matrix argument, but rather the situation is. It is easily paraphrased as *I resented that Kim mistreated my cat.* Grammatically, the differences between the two types are clear. Where *Kim* is a matrix object, no genitive is allowed, but passivisation is acceptable when the intervening NP is a matrix object. Huddleston calls *Kim* rather a subordinate subject with a matrix verb such as *resent*. In this case, the genitive is allowed, but passivisation is impossible.

(12)   a. I caught <u>Kim/∗Kim's</u> mistreating my cat. [matrix object; no genitive]
        b. I resented <u>Kim/Kim's</u> mistreating my cat. [subordinate subject; genitive OK]

(13)   a. Kim was caught mistreating my cat.
        b. ∗Kim was resented mistreating my cat.

In terms of the upcoming Chapter 6, it looks as if the complement of *resent* is an imperfect nominal with a propositional denotation. Certainly, the paraphrase given in (11) is exactly what one would expect if the complement were this type of nominal. Also note that not only are both *Poss-ing* and *Acc-ing* types allowed as a complement of *resent*, but so are the *PRO-ing* type. Again, where a verb like *catch* requires a matrix object, with the *-ing* form intervening NP is necessary.

(14)   a. I resented having to wake up early.
        b. *I caught mistreating my cat.
        c. As I emerged from the trance, I caught myself mistreating the cat.

---

[38]All examples with *catch* and *regret* as well as the table in this subsection taken from (Huddleston, 2002a).

Using (11)(c) as a guide, *see* is a hybrid of both of the above types – the intervening NP is a matrix object but not matrix argument

(15)   a. I saw Kim/*Kim's mistreating my cat. [no genitive allowed]
          b. Kim was seen mistreating my cat. [Matrix passive OK]

But, Kim is not the argument of *see*, rather the activity is. Compare this to (11)(a), where there it seems that in order to catch Kim in the act, I must catch Kim. While this also seems to be the case in (11)(c), where I most likely will also have seen Kim, this needn't always be.

(16)   a. I saw the wind rustling the leaves.
          b. I saw Kim as she was mistreating my cat.
          c. ??I saw the wind as it was rustling the leaves.[39]

It seems as if there is a syntax/argument structure mismatch, as the table below shows.

	Matrix Argument	Matrix Object
I caught Kim mistreating my cat	Yes	Yes
I saw Kim mistreating my cat.	No	Yes
I resented Kim mistreating my cat	No	No

Another uniqueness of the perception verbs is that unlike *catch* or *regret*, *saw* can also take a bare infinitive complement.

(17)   a. I saw Kim steal my cat.
          b. *I caught Kim steal my cat.

In light of ordinary language usage, this is rather unsurprising. We can easily talk of seeing an event, seeing something *happen*. In the case of (17-a), what happened is my cat was stolen – this is the implication of (17-a). But, if I catch Kim *stealing* my cat, the implication vanishes, and one is more likely to conclude the activity portion of *steal* must hold at the time of catching, and was thwarted before culminating. Now if the implication of (17-a) is because the culmination is reached, (17-b) is the exact opposite. Something is caught happening. In more technical terms, should the culmination have been reached (indicated by use of the simple form), *caught* would no longer be contemporaneous with the activity

---

[39]The acceptable form, *I saw the wind rustling the leaves* shows the interpretative aspect of activity perception. Shown a silent film of a trees in the forest with their leaves rustling, the former is a perfectly reasonable perception. From seeing the activity along with interpretation, we know that causally, *something* is rustling the leaves, and world knowledge gives us the rest. Other descriptions of this activity follow *rustle*'s taking part in the causative/inchoative distinction: *I saw the leaves rustling* or *I saw rustling*.

of *stealing*.

Interestingly, a IPVC such as (17-a) does differ from a PPVC such as (15) in that it cannot passivise. For example, there is an ungrammaticality with

(18)   a. I saw Kim steal my cat.
        b. *Kim was seen steal my cat.

Nevertheless, some authors such as (Akmajian, 1977) held that at deep structure *Kim* is a direct object of the matrix verb. However, the slightly different structures of the PPVC's and IPVC's make it that passivisation is blocked for IPVC's. It is exactly this mismatch, along with a strong semantic intuition that leads (Gee, 1977) to voice his response to the treatment of perception verbs in (Akmajian, 1977). It is the semantic notion that the entire perception verb complement is an argument of *see* that inspired the quotation seen in the section 2.4. Ideally, one would like the syntax to reflect the semantics in a way they do in the complements of *catch* and *regret*.

Already, an aspectual dimension can be gleaned from the comparison of the two examples. Arguably, part of the difference between (18-a) and (b) are the matrix verb's differing aspectual natures. Indeed, the aspectual nature of *see* (as well as *hear*, *feel*, etc.) plays a crucial role in determining their semantics. But first, a closer look at the uniqueness of perception verbs.

4.3.2   Seeing more

Aside from an apparent syntax/semantics mismatch there are tense, aspect and interpretational restrictions that taken together make perception verbs rather unique.

Tense

Compare the following

(19)   a. Mary wants to become an actress.[40]
        b. Mary regrets having become an actress.
        c. Mary saw (that) John had left.
        d. Mary saw (that) John would leave.
        e. *Mary saw John to leave.
        f. ??Mary saw John to have left.

(20)   a. *Mary saw John have left.
        b. *Mary saw John be going to leave.
        c. *Mary saw John be leaving.

---

[40](Felser, 1999) uses this example from Stowell to show that "control infinitives often express an 'unrealised' future tense." The remaining examples are adapted from (Felser, 1999).

    d. Mary saw John leave.
    e. Mary saw John leaving.

The relation of the matrix verb to the complement in (19-a) and (b) occurs in numerous catenative constructions in terms of tense. Mary's desire is in the present tense, while *become an actress* is located somewhere in the future. Sentence (19-b) has a opposite tense relation between matrix and complement. Again, the matrix verb is in the present tense, but the complement event occurred at some point in the past.

Indirect perception constructions such as (19-c) and (d) have the same possibilities. Sentence (19-c) could be a description of an inference Mary makes upon no longer seeing his car in the garage. Similarly, if she sees John a bit restless, looking at the door and playing with his car keys, (19-d) is a perfectly reasonable inference. In both cases the complements have complete temporal independence from the time specified by the matrix clause, either before or after the matrix clause.[41]

The direct perception sentences are a quite different in their temporal interpretation. If Mary saw John leave (or sees him leaving), then her perception temporally coincides with the leaving event of John. It will become clearer in next section, but with BI complements the time of the complement event is identical to the time of the matrix *see*.[42] For PPVC's (direct perception reports with *-ing* forms such as (20-e)), the requirement is less strict. If Mary sees John leaving, it is only required that he sees part of this event. For example, if Mary sees John putting on his coat and walking to the door, this perception is sufficient for her to utter (20-e), but not (20-d). Nevertheless, there is a necessarily a temporal overlap between the matrix and the complement that does hold for many other catenative verbs.

It is also important to note that a major syntactic difference between the direct perception and indirect perception sentences is the tense status of the complement. As already noted, the complement of a direct perception is non-finite, either a bare infinitive or a participial. Contrast this with the indirect perception sentences in (19). What marks them as indirect perception is not the presence of *that* or another complemetizer (this is often optional), but that the complement has morphological tense of its own. For example, sentence (19-a) is morphologically pluperfect, but could just as easily be simple past, i.e. *Mary saw that John left*. The ungrammatical sentences in (20) are examples of unsuccessful use of auxiliary tense markers in their bare form (*have* and *be*). With direct perception sentences, the complements are always non-finite and do not allow auxiliaries.

---

[41]Simultaneity is also possible, as in *Mary sees that John is leaving (and runs after him)*.

[42]This is Felser's *Simultaneity Condition* for IPVC's (direct perception reports with BI's)(p.39) which isn't quite right. Certainly, the temporal endpoint of the complement event must be included in the temporal interpretation of *see*.

Direct perception sentences do have a peculiar property that the complement contains no tense morphologically and that its semantic tense must overlap with the temporal interpretation of the matrix. The question then arises as to whether this 'simultaneity' between the time of the matrix perception verb is syntactic in nature[43] or a semantic restriction due to the nature of the perceptual activity. This will be examined in more depth in a coda to this chapter, but the reader should keep in mind a pair of possible counterexamples for now.

The first involves a variant of (19-a), where a verb of intention needn't use a *to* infinitive to capture the future tense reading of the complement. The verb *intend* can do exactly this.

(21)   a. I intend becoming undisputed champion and leaving my legacy.[44]
        b. I intend to become undisputed champion...
        c. I intend on becoming undisputed champion.

The first example is typical of non-North American English dialects,[45] while the either form seen in sentences (b) and (c) would be preferred in American dialects. Interestingly, sentence (a) seems to violate both constraints in generative grammar as well in as some versions of functional grammar, but this will have to wait until the final section.

A second counterexample has to do with the use of *see*, not in a direct perception context, but rather used as a prediction, for example what one would hear when going to a fortune teller.

(22)   I see you meeting a tall, dark stranger.

where the interpretation (in this context) is that the addressee *will* meet a tall, dark stranger. This is not such a surprise, actually. Simple present tense sentences do have a futurate interpretation, despite simple present or present progressive marking, and in these two counterexamples, something similar may be occurring.

Aspect

There is an opposition of entailment properties between IPVC's and PPVC's, whose recognition goes back several decades at least. Observe the following pair (from (Valin & La Polla, 1997)).

(23)   a. Kim saw Sandy leaving early (and stopped her and asked her to stay

---

[43]i.e., if the complement is non-finite and carries no tense marking itself, it will inherit the tense of the matrix verb. While this certainly does happen in direct perception reports, it is not always the case with other verbs, even perception verbs used in a non-direct context.

[44]www.boxing-monthly.co.uk/content/0402/three.htm

[45]The UK, Ireland, Australia, New Zealand and South Africa, at least. In many dialects of these countries both the bare complement and complement with a preposition are possible.

a few minutes longer).
b. Kim saw Sandy leave early (and called her and asked her to come
back/*and stopped her and asked her to stay a few minutes longer).

As can be seen in the follow-ups, when the complement is a PPVC, there is
no entailment of completion. Sandy never manages to leave in sentence (23-a),
but does in (23-b). The BI version does have an entailment of completion, as
can be seen when (b) is followed up by ("and stopped her..."). Compare these
entailments to those in the standard contrast between simple and progressive
aspect.

(24)   a. Sandy was leaving early, but Kim stopped her...
         b. Sandy left early, but Kim called her at home.../*but Kim stopped
         her...

Furthermore, the fact that the simple present tense for non-states is not inter-
pretable as a semantically real present tense, but rather narrative, extends to
perception sentences.[46]

(25)   a. #A man walks in the park.
         b. A man is walking in the park.
         c. I see a man walking in the park.
         d. #I see a man walk in the park.

The progressive paradox also extends beyond the complement to the whole per-
ception report itself. Recall that a standard way to differentiate an activity and
an accomplishment is by the implications of the progressive, that is

(26)   John was running in the park.

entails that

(27)   John ran in the park.

while there is no such implication from *John was running a mile* to *John ran
a mile*. Perception reports respect this as while there is a huge implicational
difference between

(28)   a. I saw John run a mile.
         b. I saw John running a mile.

there is no implicational difference[47] with

---

[46]'#' is used to denote infelicity for a real present tense interpretation. All of the examples
are perfectly acceptable in a narrative (or similar) context.
[47]Though perhaps there are stylistic and other more subtle differences.

(29)   a. I saw John run in the park.
        b. I saw John running in the park.

which is to be expected as simple activities are arbitrarily divisible. If John was running from 1 to 2, he also ran from 1 to 1:30, and perception reports follow likewise. On the contrary, if John was running a 10K run from 1 to 2, it is not the case that he ran the 10K from 1 to 1:30. Finally, (Felser, 1999) points out that when certain verbs (often achievements) are infelicitous in the progressive they are also infelicitous in IPVC's.

(30)   a. ?*She was finding her long-lost ring.
        b. ?*She was recognising her old friend.
        c. ?*I saw her finding her long-lost ring.
        d. ?*I saw her recognising her old friend.[48]

(Fillmore, 1963) is apparently the first author to probe the connections between PPVC's and the progressive.[49] In the transformational spirit of the early 1960's, he proposed that PPVC's are derived from progressives by *BE Deletion*. Transformational syntax may be long deceased, but Fillmore's basic idea was taken up again in the 1980's and has made its way into many accounts of PPVC's.

Aktionsart restrictions

(Felser, 1999) relies on the distinction in (Carlson, 1980) between stage-level and individual-level predicates. It seems similar to the structural/phenomenal distinction to arrive at a restriction on the aktionsart of the perception verb complement. Some states, such as *be obnoxious*[50] can apply to stages of individuals, and this is manifest in that such predicates readily occur in the progressive.

(31)   John is being obnoxious.

Others, such as *be intelligent*, apply only to the individual and do not appear in the progressive, being interpreted as inherent properties of an individual.

(32)   *John is being intelligent.

Felser uses this distinction to give the SLP[51] constraint which says that "for a direct perception interpretation to be available, the bare infinitive complement must contain a stage-level predicate." (p.45)

---

[48]The judgements are from Felser. Personally, I find *recognise* to be fine in both the simple progressive and PPVC, but this does not dilute the point.

[49]IPVC's are parallel to simple aspect, which in English, is completive/perfective for non-state predicates.

[50]As seen in Chapter 2, predicates such as *be obnoxious* can be interpreted as temporary behaviour of an individual or the inherent way an individual is.

[51]Stage level predicate.

Thus, when comparing an stage-level predicate (*drunk*), with the individual level (*intelligent*) the following contrast occurs

(33)   a. We saw John drunk.
        b. *We saw John intelligent.[52]

While, out of context, (33-b) is quite unacceptable, (Higginbotham, 2005) notes that if Carol is intelligent on alternate Wednesdays by taking a pill, then *I have seen Carol intelligent* is perfectly fine.[53]

In Chapter 2, it was demonstrated that a state could be coerced into having an activity reading by the use of the a copular auxiliary. This can also happen with direct perception sentences. In arguing over the merits of a new talk-show host, the following discourse between A and B occurs.

(34)   A. With those kind of shows you really need an intelligent host to carry
        the thing. DL Hughley just seems kinda...dumb. He can't pull it off.
        B. I've seen him be intelligent before, though. I don't know what the
        deal is with this show.

The respondent has seen DL Hughley behave in an intelligent way before, which is how the sentence should be interpreted.

Finally, Felser notes the ungrammaticality of the following

(35)   *We saw fireman be intelligent.

While already establishing that *be intelligent* can be a complement in a direct perception sentence, the reason is slightly more complicated. In general, bare plurals have only a generic reading with individual-level[54] predicates, but can have either a generic or an existential reading with stage-level predicates. For example, (36-a) has both existential and generic readings, with (b) has only a generic reading.

(36)   a. Firemen extinguished fires.
        b. Firemen are intelligent.

Combine this with a constraint on bare plural subjects in perception complements that says they cannot be interpreted generically, then (35) will be ungrammatical as it only has a generic interpretation. But, imagine a discourse where someone

---

[52]There is a indirect version possible, i.e. *We saw John to be intelligent*, which is an inference from his behaviour to his nature.

[53]Interestingly, in this article, Higginbotham is arguing for E-positions in even stative (normally individual-level) predicates. In the next section, we shall see that it is exactly the supposed lack of an E-position in *intelligent* that Felser claims causes the 'ungrammacticality' of (33-b). In general, Higginbotham argues that when stative verbs are used in direct perception contexts they are interpreted as perceptions of activity. I am in full agreement with this.

[54]Stative individual-level predicates and activities.

holds of firemen in general that they are unintelligent, but I disagree.

(37)   A. Firemen are not intelligent.
         B. Well, I saw firemen be intelligent just last week. They made it to the
         building before it burned down and put the fire out.

Of course, the interpretation of the subject is not generic, but it is used in a context to contradict a generic statement, and is perfectly acceptable. Therefore, while I agree somewhat with Felser's restrictions, they are interpretational restrictions, rather than the restrictions on grammaticality of direct perception sentences. (Ordinarily) individual-level predicates and bare plurals are grammatical, but they must be interpreted as being activities predicated of specific individuals.

General summary of restrictions

As the syntactic details not specifically related to the status of the complement in its tense, aspectual and selection restrictions will not be examined here, the reader is referred to (Felser, 1999, Chapter 2) for an exhaustive summary of some of the more detailed syntactic properties of direct perception complements. However, we shall sum up the previous sections of the germane tense, aspect and aktionsart properties. First, there is the simultaneity constraints on perception reports, in Felser's words, "the time interval taken up by the event described by a direct perception complement includes the time interval assigned to the matrix event" (p.81). Second, we saw that the major difference between IPVC's and PPVC's is their aspect – IPVC's have completive/perfective aspect, while PPVC's have progressive aspect. In the last section we saw what Felser calls the SLP constraint – direct perception complements must contain a stage level predicate. In our terminology, the complement must be interpreted as an event (or activity). Any typically stative verb used in a perception verb complement is only interpretable as a stative verb coerced into an activity meaning. Given the aspectual restrictions on perceptual complements this is no surprise – states have neither completive or progressive aspect, though it is held here that these constraints are interpretational rather than syntactic. Felser also chooses to follow the tradition that the logical properties of perception sentences (verdicality, etc.) are non-default implications. Finally, there are a number of more syntactic constraints ruling out complementizers and finiteness markers,[55] expletive *there*, and passivisation of the intervening NP, which will not be examined here.

Status of the complement

Armed with the above constraints and properties of perception verb complements, Felser sets out to give a structure for them in minimalist syntax. The number

---

[55]Complements with either (or both) of these are indirect perception complements.

of previous analyses of PVC's are legion and the interested reader is invited to see (Felser, 1999, Chapter 2) for the relevant history.[56] Felser takes Fillmore's progressive hypothesis seriously, noticing that the *only* major difference between PPVC's and IPVC's is their aspect.[57] Both types are perception verb complements are said to be aspect phrases, i.e. maximal projections of the functional head *Aspect*.

Not being a syntactician,[58] there is very little to say on this point, or even the status of a functional head *Aspect*.[59] This head comes between **T** and *v*, meaning that both the complement VP and aspectual information is part of the aspectual phrase, but it is unspecified for tense. What is important for our purpose, is that on this account, the complement is taken to be a constituent as a whole, and has aspect, but no tense. For the more syntactically naive amongst us, the aspectual phrase has everything needed to be a simple sentence,[60] but tense.[61] Given that we want to follow Gee's initial intuition,[62] and treat PVC's as denoting events, a structure that carries both lexical verbal information and aspect is the perfect candidate.

A simplified version of an Aspectual Phrase that serves as the complement to *I see John draw(ing) a circle* can be see in the following tree.[63]

---

[56]Attempts include PVC's as IP's, bare infinitive *there* clauses, bare VP's, auxilliaries, small clauses and predicate phrases.

[57]There is one ostensible difference, however. PPVC's appear to passivize, as in *John was seen stealing my car*. Recall, that this is not always allowed, i.e. #*The wind was seen rustling the leaves*. Felser proposes that a sentence such as *I saw John stealing my car* has two structures – one of a PPVC, and one that is more similar to a complement of *catch*. In the first, it is the event or activity that is seen. In the second, it is the agent that is seen while doing an activity. When direct perception reports involve a situation where the agent cannot be seen, only the first structure is available.

[58]Or being wedded to a particular syntactic theory.

[59]As opposed to a lexicalist account or subsuming it under an inflectional head.

[60]i.e. the complement of *I saw John steal my car* is very close to *John stole my car*, but lacks any tense.

[61]This is a bit of an over-simplification, but as we are most concerned with the 'eventive' relation between matrix and complement many of the syntactic details are extraneous. But do note that as the aspectual head come with no tense specification of their own, and also come before **C**, complements and auxiliaries are also ruled out of PVC's.

[62]And taken up in the event semantics field by Higginbotham and followers arguing against a propositional interpretation of PVC's, but rather an 'eventive' interpretation.

[63]Simplified in the sense that arrows for particle movement and case checking to account for purely syntactic features that are not germane to this discussion are left out.
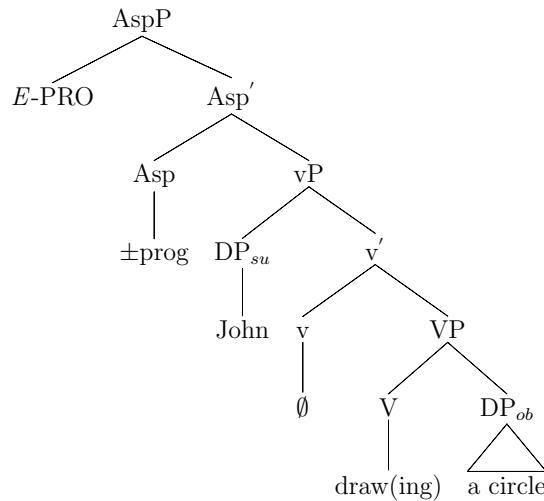
The aspectual phrase falls under the AP of the matrix verb,[64] which in turn, having tense, falls under a TP, which gives the tense. The event variable of the matrix (**E-pro**$_i$)[65] receives tense from **T** and via event control, sets the tense of **E-PRO**$_i$ in the complement AP. This can be seen, schematically, below:

$$[_{cp}C[_{TP}Spec[_{T'}\mathbf{T}_i[_{AspP}\mathbf{E\text{-}pro}_i[_{vP}...[_{AspP}\mathbf{E\text{-}PRO}_i vP]]]]]]$$

Every non-stative[66] verb is realised as an AspP – the rightmost AspP (the complement for e.g. *I saw John draw a circle*) dominates a vP (this is a DP v'[67]) which is 'John draw a circle'. More precisely it is Asp' that right-branching dominates a vP, and left branching spells out aspect (+/- prog). In turn, Asp' is dominated by AspP, which branches to the left with *E*-PRO. All non-stative verbs are realised as AP's. Inflected verbs are then dominated by a TP and CP. In the above there are two AP's – the one to the left is 'I saw' and the latter is the AP that it selects. The E's serve to bind tense.

For those who can't understand the preceding formulation, the important part to remember is that an aspectual phrase has no tense of its own, but receives it from the higher, matrix verb, which as the name suggests, does carry its own grammatical aspect.

---

[64]All stage predicates are realised as AP, syntactically.

[65]Lower case **pro** is the controller while **PRO** is the controlled.

[66]Stage-level predicate in Felser's terms. Individual level predicates are ruled out as ungrammatical as clauses projected by an Individual level predicate have no event argument, thus ruling them out as AP's.

[67]Which in turn is a v V.

4.3.3   Semantics

Event control analysis

The last section briefly stated what Felser takes a PVC to be syntactically, and helps to explain certain syntactic properties as well as the aspectual difference between IPVC's and PPVC's. However, more is needed to explain the simultaneity of tense as well as the logical properties of perception reports.[68] This is achieved by what Felser calls an event control analysis.

Briefly,[69] all stage-level predicates are realised as Aspectual Phrases. The specifier, in English, of AspP is a position which holds a stage-level predicate's external argument. This external argument is a Kratzerian event argument E. This differs from a standard Davidsonian event variable in that it is determined syntactically (not lexically) and serves semantically only to locate the spatio-temporal location of an event.[70] Being, an external argument, the event argument is realised outside of the predicate. The two conditions on Event Arguments are that they are licensed by an aspectual phrase, and furthermore they must receive a temporal index.

In simple, tensed (stage-level) sentences, this is done by **T** position, i.e. *John drew a circle* would have an aspectual phrase that is non-progressive, as well as an external event argument. This argument is then assigned a temporal index by **T**, in this case, sometime in the past. But, when there is an non-finite (untensed) aspectual phase, such as *John draw a circle* in *I saw John draw a circle*, matters become slightly more complicated. Analogous to anaphoric PRO, there is an *E*-PRO that allows an non-local event argument to receive a temporal index.[71] *E*-PRO is like the obligatorily controlled PRO that is co-indexed by the event position in the higher clause. This is Felser's *Event Control Hypothesis* (p.146)

> In direct perception constructions, the perception verb functions as a control predicate in that its event argument controls the event argument provided by the embedded predicate.

Stated more plainly with the example *I saw John draw a circle*, there is an event argument that is associated with the matrix verb (i.e. *see*) – this argument receives its temporal index from T. The IPVC is an untensed aspectual phrase whose event argument needs receive tense. As there is no **T** locally, the event argument of the IPVC is bound by the event argument of the matrix verb. Thus,

---

[68]There is also a syntactic account for the non-passivisation of perception reports. It is more than that a PVC is a complete constituent, but involves purely syntactic issues that are of no concern for us.

[69]The interested reader is directed to (Felser, 1999, Chapter 4).

[70]In a discussion, she notes some authors have attempted to reconcile the two concept to the effect that there is both a syntactic event argument in Kratzer's sense as well as a Davidsonian event variable that is lexically determined.

[71]Standard examples of PRO in action are sentences like *Bill$_i$ decided [PRO$_i$ to get up early]*.

the temporal index of the complement is identified with that of the complement, deriving simultaneity. There are two Asp's, the leftmost corresponding to [*see* Asp], and the rightmost corresponding to the complement. Note that the event argument of *see* is E-pro, while that of *John draw a circle* is the controlled E-PRO, giving them both the identical temporal index assigned by T. A simplified version, looking only at the event binding relations (and the sentence filled in for ease of reading) is

I **E-pro** saw [John **E-PRO** draw a circle.]

The important part to note are that only stage-level predicates receive a E-variable. This means that *see* must be a stage-level predicate and that individual-level predicates such as *(be) intelligent* are ruled out as complements. Also, generic sentences are only interpreted on an individual-level, thus also not allowed in direct perception constructions.

While we agree with the simultaneity condition for direct perception sentences, the way Felser derives it is primarily syntactic. IPVC and PPVC complements are aspectual phrases that have no tense of their own, and instead receive it from the matrix. But,[72] the final section of the chapter will show that doubt should be cast on this more generally. It may be only an accident of dialect rather than a hard syntactic constraint.

Event Control and Logical Properties

On the event control analysis both veridicality and substitutivity follow rather directly. On the syntactic level PVC's lack the categories **C** and **T** and do not allow sentential operators. The lack of any of these properties means that the two events are linked together. Felser writes that "direct perception statements form a single proposition in the sense that they are associated with the same (referential) T, and share a single truth value" (p.177). Thus the truth value of the matrix event implies the truth of the complement. This takes care of veridicality and, as there is no referential opaqueness, substitution as well. The only question is whether or not determining such implicational properties purely by syntactic means might be a bit too strong. The discussion at the beginning of the chapter casts doubt on such strong logical notions.[73] In ordinary language, one would deny that there were pink elephants dancing in the room, but it seems rather strange to tell the alcoholic that he didn't see them dance.

---

[72]As already hinted.

[73]In (Felser, 1999, Chapter 2), she qualifies veridicality with the caveat, "in the normal case", which is absolutely correct. But to then derive the principle by syntactic means alone makes it rather hard to determine how 'normal' fits in any longer.

Problems and unfinished business

Syntactically, this proposal seems to work and provide not only simultaneity, but also the various logical properties such as veridicality, substitution and exportablity. However, there are a few issues that need to be addressed. The first is that her arguments for *see* being a stage-level predicate are not that convincing.[74] However, we do concur that *see* is not an ordinary state or individual-level predicate. To do this one must take a closer look the event-type for *see*, which is taken up in a later section. The use of a very syntactic event argument by Felser does justify her not taking such a close look at the semantics below the event variable level; but the unsatisfing arguments that *see* as a stage-level predicate make a closer look at the aktionsart of *see* to be a worthwhile enterprise if we want to take it as an event. The second issue arises when one looks at what happens to the progressive hypothesis when *watch* is used in place of *see*.

See as stage-level predicate

Felser's proposal is that *see* is a stage-level predicate and has an event variable in its thematic grid. The complement has completive or progressive aspect, but no tense. Thus it has an event variable, but its tense still needs to be anchored – there is no TP layer in an aspectual phrase (and these are Kratzerian event arguments that need to receive a temporal index). The event control works very much like PRO in the argument world.

This proposal is enough to get the tense simultaneity, and is compatible with the durative type of complement with a progressive, the perfective and possible instantaneous type with a BI. However, there are two problems. I agree that *see* is not a state, but the arguments that it is stage level in contrast to supposed individual-level only verbs do not convince. For example, appearing in the progressive is a hallmark of a stage-level predicate. But she gives examples of *see* with PPVC's are either non-veridical or not an event of direct perception. Thus, whether direct perception *see* is a stage-level predicate or not cannot be concluded from them.

(38)   a. John thought he was hearing those voices again.
        b. I am hearing lectures.
        c. He was seeing them robbed day by day.

The last two are rather iterations of many *seeing* events. However, there are perfectly good examples of *seeing* as a single, perception event. However, in order to make her point that *see* is stage level, she contrasts it with *know* – the paradigm of an individual predicate. *Know* is fine in the progressive, as is detailed in the introduction to this thesis.

---

[74]*See* must be a stage-level predicate in order for to contain an event variable, allow for event control and tense binding.

She also says that true statives such as *know* cannot be embedded under the verb *force*, while *see* obviously can be. So

(39)   *Mary forced him to know the answer.

This does sound bad, but as I presume *believe* must be just as much of a state as *know*, it is no stretch to say "I was forced to believe that...", by my parents or the church. In any case, "forced to know" is quite common in ordinary usage.

(40)   a. So they forced me to lose my job and *they forced me to know that they did it to me on purpose*, and that it was related to the things that the same family has done to me for about 40 years.
        b. Having a neighbour like this has been an adventure, to say the least. *He has forced me to know the local bylaws more than I ever wanted to.*

There are many such examples, and arguably they could mean the end-state of learning something and be a sort of coercion. However, this is not the case with *believe* and we can find many examples like the following.

(41)   So my search for truth forced me to believe that language is not an exact science.[75]

She also argues that *see* is fine in an imperative, but individual-level predicates are not. We can say, "See that man over there", but not "Know the answer" or "Be tall". However, *Know the answer, dammit* is perfectly fine to tell a quiz-show contestant at the appropriate moment, and any legal-aid foundation will tell you to 'Know your rights!'. Thus, I can see little justification from these cases for treating *see* different from verbs such as *know*.

Nevertheless, there is certainly a huge event type difference between *see* and typical states that belies it use in the simple present. But this is also a problem, as will be shown in section 4 of this chapter, in that there is more than one event type for *see*;[76] this depends on the event type of the complement. Felser never goes below the event variable and pays little attention to Aktionsart except to ban individual-level predicates. But if the event variable of the complement is controlled by that of the matrix, how does the event type of the complement determine that of the matrix? Moreover, there is no explanation as to what happens when the complement is morphologically simple, but must receive a progressive interpretation. To be fair, this is a property not of *see*, but it's activity counterpart.

---

[75]www.beneficentchurch.org/srmn101903.html
[76]This is done through the methodology discussed in Chapter 2. Aktionsart is often not a property of a verb or VP alone, but rather the entire sentence in context.

The peculiarities of *Watch*

Felser says very little about *watch*, but as a preview of the next section, direct perception sentences with *watch* in the progressive cause their complement to also be progressively interpreted, even if the complement is an IPVC and presumably of completive aspect.

(42)   a. I was watching Mary cross the street when she was hit by a car.
        b. I watched Mary crossing the street when she was hit by a car.
        c. ??I watched Mary cross the street when she was hit by a car.

Unlike *see*,[77] the aspect of the matrix verb controls not only the tense of the complement, but also the aspect. If the complement in (42-a) is an aspectual phrase that comes with its own aspect, it brings up the question as to where the progressive interpretation for a simple form aspectual phrase comes from.

4.4   Preformal matters

Having established that the eventive complements of perception verbs are semantically events, that have their own aspect,[78] but are derivative on the matrix clause for tense, we can begin examining further properties that are necessary for a formalisation. At the beginning of the chapter we saw a tension between a perception statement of direct experience and a one where interpretation is part of the report. The examples where the speaker reports an other's presumed perception (by inference) or there own at a later date, once new information is known are perhaps the most interesting and are examined further in the next section. Following that, as it is necessary to first understand the aktionsart of *see*, a section will be devoted to that. Finally, a contrast will be made with *watch*.

4.4.1   Aktionsart of See

Felser needed *see* to be an individual level predicate in order for there to be a matrix event, and thus control of the tense of the complement. While her arguments are somewhat wanting, it is certainly the case that *see* is no ordinary state, despite the present tense appearing in the simple form as a default. Instead, the part of perception that is direct and experiential renders *see*, in a sense, transparent to the event that is described. This can be demonstrated by examining the

----

[77]For *see* in simple aspect. When something like *I'm seeing Mary cross the street* is acceptable, the complement is also interpreted as progressive. I have no explanation, but somehow when the matrix is in the progressive aspect, it determines the aspect of the complement.

[78]This last part holds for *experiencer* perception verbs such as *see* or *hear*. When the more agentive versions are examined (i.e. *watch* or *listen*), the relation of aspect between matrix and complement becomes more complicated.

standard behavioural tests for Aktionsarten (see Chapter 2), but first, a lovely impressionistic account of perception verbs taken from (Huddleston, 2002a).

> The non-modal construction[79] is at the boundary between stative and dynamic. *I heard a plane pass overhead* contrasts clearly with *I could hear planes passing overhead* as dynamic (an achievement) vs. stative, but we can also have *I heard the tap dripping*, which is also state-like, differing little from *I could hear the tap dripping*. In the present tense, the simple form tends to sound somewhat more dramatic, suggestion a quasi-dynamic interpretation: *Yes, I see it now; I smell something burning.*

While terms such as 'state-like'[80] and 'quasi-dynamic' may seem a linguist's equivalent of a the description on a label of a bottle of wine, it is no accident. The different events described are normally construed as having a certain structure. A plane that passes overhead (aurally) is a quick bit of noise, altered by the Doppler effect, and then quickly disappearing. Iterated as a series of plane-passing events, the quality changes to focus less on a quick change, but rather a series of events envisaged as a continuing activity. Perception verbs unsurprising reflect these distinctions. But not content at leaving the matter at the poetic level, we shall see what happens to the event type of *see* given different Aktionsart and aspect configurations of the complement.

Seeing an activity or accomplishment

Recall from Chapter 2, that an activity or accomplishment is only felicitous in (real) present tense in the progressive aspect.[81] The simple form can be used in the past tense, but this implies that the event is no longer occurring. This question remains open if past progressive is used. Thus

(43) a. #Mary jogs in/across the park.
   b. Mary is jogging in/through the park.
   c. #Mary just jogged in/across the park. She still is.
   d. Mary was just jogging in/through the park. Maybe she still is.

---

[79]We have not examined these sorts of modal constructions, but they have a less direct, more neutral 'flavour'.

[80]Huddleston seems to use the term 'stative' for iterative activities (i.e. the tap dripping, planes passing overhead). Certainly such progressives profiling uniform activities have a more 'stative' quality than accomplishments or achievements, most likely due to the fact that any part of a 'tap-dripping' event will be much the same as any other. Nevertheless, they are not states.

[81]Excluding narrative and habitual/generic uses.

See can be used with the above complements in the simple present tense, *if* the complement is in the progressive, but not if the complement has simple aspect.[82]

(44)   a. #I see Mary jog in/through the park.
          b. I see Mary jogging in/through the park.
          c. I just saw Mary jogging in/through the park. #I still do. Maybe she still is.
          d. I just saw Mary jog in/through the park. #I still do. #Maybe she still is.

The examples in (44) show an interesting interaction between the aspect of the complement and the aspect of the entire sentences and/or discourse. In sentences (44-a) and (b), the aspect of the complement determines the felicity of the matrix verb in its tense/aspect configuration. Here *see* behaves like an ordinary dynamic event, if we assume that the aspectual status of the complement is mirrored in the matrix.

But sentence (44-c) complicates matters considerably. As evidenced in the continuations, a progressive complement allows continuation of the event into the present. However, *see* now behaves as the simple form it actually is. If I saw Mary *jogging* in the park a minute ago (and express this using a progressive complement),[83] it is certainly possible she still is. But the *seeing* part is completed, regardless. With the simple aspect complement in (44-d), it is not only the *seeing* that is finished, but, from the speaker's point of view, also the jogging activity. In reality, what the speaker saw as and end to the activity, may have merely been a stretching break. But the activity is *presented* by the speaker as complete, thus making the continuation infelicitous.

For present tense, *see* behaves like its dynamic complement, only felicitous when its complement is also felicitous in simple present. In past tense, it is stronger, and always behaves like a past tense simple form of a dynamic event type. This is more reason to think that *see* is not a typical state.[84]

---

[82]Given that the simple aspect complement is perfective or completive, the object of present tense *see* has an extended temporal profile. Given the semantics of *see*, it would then have the same (or similar, allowing that all the event needn't be seen) temporal profile. Therefore, the reasons given in Chapter 2 for the infelicity of simple aspect activities and accomplish in real present tense hold for *see* as well.

[83]The notion of how the speaker expresses the situation is quite important. Should I have used the simple form, I imply that the event is completed (even for an activity). Mary may still be jogging, but this cannot be recovered from the way I construed and expressed the situation.

[84]We have yet to see an example of *see* in the progressive. The reason is that I have very little understanding of this sort of usage that is not iterative or habitual. However, *watch* appears quite normally in the progressive aspect. We will discover that the aspectual interactions between matrix and complement become even more complicated when a progressive matrix is brought into the picture.

Seeing a point

Similar to dynamic event types, *points* are also infelicitous in a simple present configuration. When used in the progressive, the are coerced into an iterative activity, construing a series of points into a uniform activity. In the past tense, simple form can be used, and can mean either that one instance of the point happened, or an iterative activity happened.

(45)   a. #The light flashes.
       b. The light flashed. (once)
       c. The light flashed several times.
       d. The light is/was flashing.
       (iteration of flashes)

Embedding these complements under *see*, the pattern remains the same.

(46)   a. #I see the light flash.
       b. I saw the light flash. (once)
       c. I saw the light flash several times. (iteration of *seeing* and *flashing*)
       d. I see/saw the light flashing.
       e. I keep seeing the light flash.

Should the event be punctual, the experience of it will be as well,[85] making *seeing a point* having the aktionsart of a point. (46-d) behaves just as *see* with a progressive activity complement, which is as to be expected. Sentence (46-e) with the aspectual verb is a bit more interesting. Though the complement is in simple aspect, a point interpretation of it is impossible. The dynamicity of the matrix coerces the complement into having an iterative reading.

Seeing an achievement

Achievements are similar to points in that they are instantaneous (though with an added change of state). The difference comes when they are used in the progressive. They do not become an iterated activity, but something resembling an accomplishment.

(47)   a. #Mary reaches the summit.
       b. Mary reached the summit.
       c. Mary is/was reaching the summit.

Sentence (47-c) does not mean that Mary is continually getting to the summit, but rather that she is in the middle of a preparation activity (e.g. *climbing*), that, if successful, will lead to the achievement. While a run-up activity can also be inferred in (47-b), the sentence profiles only the instantaneous change of state

---

[85]Ignoring examples of extreme slow motion film and the like.

that is the hallmark of an achievement. These properties are parallelled under embedding with *see*.

(48)   a.#I see Mary reach the summit.
         b. I saw Mary reach the summit.
         c. I see/saw Mary reaching the summit.

For sentence (48-b) to be felicitous, it is enough that I see only the instantaneous change that brings Mary from a non-summit state to a being-on-the-summit state. With an accomplishment complement, such as *I saw Mary cross the street*, it is a rather odd thing to say if I only saw the last step Mary took, rather than at least some portion of the activity sub-event as well. Sentence (48-c) is the perception of a run-up achievement (i.e. an achievement coerced into a progressive accomplishment structure) and has the same properties as seeing an accomplishment.

4.4.2   *Watch* and aspect of the complement

Aktionsart aside, an interesting phenomenon happens when *see* is replaced by *watch*. Felser has little to say about *watch* aside from the observation that it cannot be used with either indirect perception complements or infinitival complements.

(49)   a. #We watched that he drew a circle.
         b. *We watched him to be an impostor.

Watch is a more intentional verb than *see*, and it is also an unabashed activity. Unlike *see*, but typical of activities, the simple, present form of *watch* can only be interpreted as a narrative tense. For ordinary present, the progressive must be used.

(50)   a. I watch John crossing the street. (narrative only)
         b. I'm watching John cross the street.

Presumably the complements of *watch* are the same sorts of bound aspectual phrases that are associated with *see*.[86] If this is so, then something interesting happens with the aspect of the complement that is then left unexplained.[87] The complement of (50-b) is a non-progressive aspectual phrase. As expected, there is temporal simultaneity between the matrix and complement, but there is also a change of aspect. (50-b) in no way implies that John crosses the street. Compare

(51)   a. I was watching John cross the street when he was hit by a car.

---

[86]Chapter 4 of (Felser, 1999) does extend the event control analysis to apply to aspectual verbs as well.

[87]In cases where *see* can be used in the progressive with an activity, this will also be the case.

　　　b. I watched John crossing the street when he was hit by a car.
　　　c. I watched John cross the street #when he was hit by a car.

When *watch* is in the progressive, the complement must interpreted likewise, even when the complement is morphologically in simple aspect. Here it seems that not only is the temporal index controlled by the matrix verb, but also the aspect. I have no syntactic explanation for this, but it is not particularly surprising given the agentive and attentional nature of *watch*. Sentence (50-b) is interesting in that while there is no strict implication of completion, the speaker is approaching what he sees in that way – he assumes that John will make it across. Yet this is the event he is in the middle of watching.[88]

When *watch* is used in the progressive (obligatory for present tense use), the aspect of the complement is also progressive (semantically), whether it is morphologically a bare infinitive or an *-ing* form. For example, suppose *I am watching John cross the street* is uttered at time T.[89] At time T, it also holds that *John is crossing the street*. Moreover, as with (51-a), the default that John actually crosses the street can easily be cancelled.

There is also a double telicity when *watch* is used in the progressive. The previous examples show that the complement event needn't complete, but, as expected *watch* carries its own telicity.

(52)　　As, I had to wait for her anyway, I was watching Mary cross the street to meet me. I quickly bored of it, and looked at advertisements until she reached me.

Here the complement did (for Mary's sake) happily complete. But the *watch* event did not. In this example, while I did watch her *crossing* the street, I preferred to whet my consumerist appetite rather than watch her *cross* the street. With this in place, we can now examine the Aktionsart of *watch*.

Watching an activity/accomplishment

As noted already, *watch* always needs to be in the progressive in (real) present tense, as with an activity or accomplishment. But unlike *see*, the standard alternation between simple and progressive forms patterns has implications on the telicity of the entire sentence.

(53)　　a. #I watch Mary cross/crossing the street.
　　　b. I watched Mary cross the street.

---

[88]Overheard on the radio before the Six Nations France-England rugby match, in an interview with supporters at an English airport: *We're going to watch England beat France. Hopefully.* The continuation could refer to them getting there to watch the match, but more likely is their prediction of the outcome. Even if they did make it to the match, they were sorely disappointed.

[89]To simplify matters, we assume ordinary veridicality of the complement for this example.

       c. I was watching Mary run in the park.
       d. I was watching Mary cross the street.

(53-b) implies telicity on the part of both the events matrix and complement. When the complement is an accomplishment, so is the entire sentence. When the complement is an activity (as in (53-c)), the entire sentence is as well. Sentence (53-c) implies that it is also the case that

(54)   I watched Mary run in the park.

As noted earlier, with sentence (53-d) there is telicity normally associated with a VP *cross the street*, but now also with the matrix verb as well. There is no implication from (53-d) that I *watched* her cross the street. It then is safe to say that watching an activity or an accomplishment is just the event-type of the accomplishment. The major difference between *see* and *watch* is that the aspect of the complement has much less affect on the sentence as a whole. The determining factor is whether or not *watch* is in simple or progressive aspect.

Watching a point

Again, the ability of *watch* to more readily appear in the progressive has both aspectual and event-type coercion effects upon the complement.

(55)   a. #I watch the light flash.
         b. I'm watching the light flash.
         c. I watched the light flash. (once)
         d. I watched the light flash 5 times.
         e. I watched the light flashing.
         f. I was watching the light flash.

When *watch* is in the progressive, a present tense *watch* report is perfectly fine in present tense. The same holds for *The light is flashing*. But,[90] as sentence (55-b) coerces the complement into having a progressive reading, the complement is forced into an activity reading. The entire sentence (55-b), is then an activity.

    Example (55-c) can be seen to be a point. Note that the progressive versions in (55-b) and (f) has a point coerced to an activity reading, as is standard with points. Progressive aspect can be on either the matrix verb or complement and seems to give the same reading.[91]

    Example (55-c) seems to be a activity constructed out a specific number of semelfactive events. This is now in the territory of the quote from Huddleston, and only impressionistic comments are possible. But the report of (55-d) has a

---

[90]Ignoring a futurate reading.

[91]Something like, *I'm watching the light flashing* also has the same reading, but is stylistically awful and unnecessary, especially as the double *-ing* has added aspectual impact.

less uniform activity feeling that the progressive version in (55-f), even if they are describing the same situation.

Watching an achievement

The last case to be looked at is the interaction of *watch* with an achievement complement.

(56)   a. #I watch Mary reach the summit.
        b. I watched Mary reach the summit.
        c. I'm watching Mary reach the summit.
        d. I watched Mary reaching the summit, but she slipped at the last foothold.
        e. I was watching Mary reach the summit when she slipped at the last foothold.

At first glance, this appears to be quite similar to *see*.[92] However, sentence (56-b) does pick quite a major difference between the two verbs. One can see an event by only seeing a split second of it. For example, if shown a stop-motion picture of a runner, the body position and context is enough to say, "I see him running". That is, the temporal profile of *see* can be an instant or a drawn out event. Recall that *I saw Mary reach the summit* can easily profile only the instantaneous change of state, and ignore the run-up activity.

But with (56-b), it is far more likely that the time profiled also includes at least some of the climbing event as well. With *watch* an instantaneous temporal profile seems impossible, and the matrix becomes a simple past tense accomplishment. Again, the progressive *watch* coerces the complement to have a progressive reading, in this case a run-up achievement. The complement may also be in the progressive, also giving the complement run-up achievement reading. As with *watch* plus accomplishments with a progressive reading, the aktionsart of *watch* becomes an accomplishment itself. Sentence (56-c) transformed into a past progressive could easily be continued with, *But I got bored and went back to cabin.*

4.5   Formalisation of *see* and *watch*

In this section we will 'see' a formalisation of the tense, aspectual and implicational properties discussed in the last section. Before moving on the nuts and bolts, a few remarks are in order. The first is that we have a constraint on the 'seeing', which we are calling *ready*, as shorthand for perceptual readiness. This is a bit of a 'catch-all' constraint, as it is not just 'looking at'.[93] Rather, as discussed in the second section of this chapter, one can look at something and

---

[92]Ignoring, for a moment, the progressive in (56-c).

[93]Or for *hear*, whatever the equivalent expression would be for 'having one's ears pointed'.

not see it, for esoteric reasons of not 'knowing how' to see an object or event, or mundane reasons of inattention, distraction and the like. In a sense, we could say that *ready* is shorthand for 'ready, willing and able.'

4.5.1   Seeing a point or achievement

In Section 4, there were a number of restrictions on perception verbs with instantaneous events discussed. As with sentences involving instantaneous events, the perception analogues are only felicitous in the simple past.[94] The entire event described by the sentence is instantaneous, giving *see* in this context a point-like temporal profile. The basic ingredients are listed below, with two punctual events, and the fluent *ready* as described above, and defined as follows:

1. The verb *See* is represented as a canonical punctual event see_e

2. e_c is the canonical punctual event of the complement.[95]

3. as a precondition on perception, a fluent *ready* is required.

4. see_e is defined in relation to its complement event: *HoldsAt(ready,t)* $\wedge$ *Happens(e_c,t)* $\rightarrow$ *Happens(see_e,t)*

As *see* carries the morphological tense (e.g. *I saw the light flash*), it is this event that is subject to the tense integrity constraint:

The query ?*Happens(see_e,R)*, $R < now$ `succeeds`, and the query ?*HoldsAt(f_1,R)*, . . . , *HoldsAt(f_n,R)*, $R \geq now$, *Happens(see_e,R)* `fails`.

The definition of *see* ensures that the *seeing* and the complement event happen at the same time and are instantaneous. The *ready* fluent also must hold at the same time, but, being a fluent and not a canonical event, it may have been holding for a arbitrary time before. For example, I could be eagerly waiting for a light to flash. The integrity constraint for the tense requires the reference time to be in the past.

The other issue is that of veridicality. Note that the integrity constraint only enforces that *see* happens. When starting the derivation with *Happens(see_e,t)*, the database then looks to verify the conjuncts in the antecedent. For example, if the sentence was *I saw the light flash*, e_c is the canonical event of a light flash. The database will then look for this event happening at t to verify the scenario. But, if it was not a flash of a light, but an atmospheric phenomenon that happened (at t), the database will unify *Happens(e_c, t)* with this.[96] Thus, veridicality is a

---

[94]Excluding narrative contexts.

[95]In the formalisation of the various aspects of *see*, the subscript c is used to indicate that the event of fluent belongs to the complement event scenario.

[96]The more complicated example of *stealing/borrowing* will be handled at the end of the section.

strong default, but does allow mistakes.[97]

4.5.2   Seeing an activity or accomplishment

Matters are a bit more complicated here. First there is the implicational difference between a progressive and simple aspect complement. Also, *see* can appear in either simple present or past tense.[98] Finally, unlike *seeing a point*, the *see* event takes time, but the temporal profile also differs depending on whether the complement is in simple or progressive aspect. There are three possible cases here.[99] With present tense *see* only a progressive complement is allowed, but with *saw*, both simple and progressive aspect are allowed in the complement.

Present tense with progressive complement

In this case, the telicity of the complement event is irrelevant. *I see Mary crossing the street* and *I see Mary running in the park* have the same lack of implications. What does need to be taken into account, however, is that *see*, takes time and has a special kind of dynamics, in that it is driven by the complement event in, to use Huddleston's words, a quasi-dynamic way. To connect *see* with the complement and obtain the proper dynamics, the following scenario is used.

1. ?*Terminates*$(e,ready,t)$, ¬*Terminates*(e,see,t) `fails`

2. *Releases*$(start_c,see,t)$

3. *HoldsAt*$(ready,t) \rightarrow$ *Trajectory*(f_c,t,see,d)

The above constraints ensure that if an event should terminate my readiness (e.g. being instantly struck with blindness, or getting bored and turning away), the seeing shall also stop.[100] The second line of the scenario allows the dynamics of *see* to be started. It it the *start* event of the complement event that does this. The third line is the dynamics proper, and says that as long as readiness holds, *see* is driven by the activity fluent associated with the activity or accomplishment complement.[101]

The template above integrates the complement event in the sense that (e.g. for *crossing the street*) it is the start of the complement event that releases the

---

[97]In the limiting case, the complement event could be unified with a hallucination.

[98]The progressive *see* works analogously to *watch*, which will be treated formally in the next section.

[99]Plus slight telicity differences depending on whether the complement is an accomplishment or an activity.

[100]The literal reading of the constraint is that no event shall terminate *ready* and not terminate seeing.

[101]As well as points or achievements coerced into activities/accomplishments, such as, *I see the light flashing* or *I see her reaching the top*.

dynamics and the *cross* fluent that drives the trajectory. For the sake of clarity, the last two lines of the above scenario for the specific *I see Mary crossing the street* are as follows:

1. $Releases(start_{cross}, see, t)$

2. $HoldsAt(ready, t) \rightarrow Trajectory(cross, t, see, d)$

The only remaining ingredient is the tense integrity constraint for *see*. As it is 'quasi-dynamic', it is treated *as if* it were a progressive.

Integrity constraint: $?HoldsAt(see, R)$, $R = now$ `succeeds`

Again, there is a default veridicality as the integrity constraint only requires that *see* succeeds, while the complement needs to emerge from the derivation and could be unified with another event if need be. The temporal simultaneity is assured as it is the activity fluent of the complement that drives the *see* event. The derivation starts with $?HoldsAt(see, R)$, which then reduces to a query of the dynamics, and as the *trajectory* predicate includes the complement event, this is then derived. Finally, as only the *start* event and activity portion need to be accounted for, there is no implication as to whether the complement stops or completes.[102]

Past tense with simple aspect complement

This is the most complicated case, and encompasses both accomplishment complements (*I saw Mary cross the street*) and activity complements (*I saw Mary running*), with the major difference being that the accomplishment does complete successfully, while the activity is presented (by the speaker) as having ended, but in a more arbitrary way. We shall focus on the more complicated, former case. Recall from the last section that *I saw Mary cross the street* implies both that Mary made it across and I saw it until the end. Importantly, however, I needn't see the event from beginning until the end,[103] but could have seen only the last bit of the *crossing* activity along with the final bit where she makes it to the other side.

As the event is one of temporal duration, the *see* event also has this profile, and, being past tense, will have be constructed via hierarchical planning in the normal way.[104]

---

[102]Depending on whether the complement event is an activity or accomplishment.

[103]Though this is the limiting temporal possibility.

[104]Though the $finish$ event can be omitted.

$Happens(start,$s$) \land Happens(finish,$r$) \land$ s $<$ t $\leq$ r $\land HoldsAt(see,$t$) \rightarrow Happens(see,t).$

What remains to be done is to connect the conjuncts of this formula to the complement event. The $HoldsAt$ conjunct is already connected via the dynamics for *see* given in the last section. The *start* event is a bit more complicated. Recall that for *I saw Mary cross the street* I can see it from the beginning until the end (assuming I am *ready*), or from a point in the middle of the complement event until the end. Thus, $start_{see}$ is defined by the following two clauses:[105]

$Happens(\text{start}_c,t) \land HoldsAt(ready,t) \rightarrow Happens(start_{see},t)$
$HoldsAt(f_c,t) \land Happens(e_{start},t) \rightarrow Happens(start_{see},t)$

where e_{start} is any event that starts *ready*, for example, suddenly looking in the direction of the street. The first disjunct is where the seeing starts from the beginning of the event, and the second disjunct is where something triggers readiness, and seeing begins.

Finally, a constraint is needed to ensure that the seeing makes it until the end. This can be done by defining a $finish_{see}$ event in two clauses, that (for telic events) synchronises the $finish$ event of the complement to be the $finish$ event of our definition of *see*. Second, *ready*, and therefore seeing, must still hold.

$Happens(finish_{see},t) \rightarrow Happens(finish_c,t)$[106]
$Happens(finish_{see},t) \rightarrow HoldsAt(ready,t)$

For accomplishments like *cross the street*, this means that the crossing is seen until the end.[107] For activities like *run*, as the atelic activity can be divided arbitrarily, it is presented as complete,[108] though, in reality, may continue arbitrarily in the future.

The only remaining thing to do is to specify the integrity constraint for *saw an activity/accomplishment* is to formulate the appropriate integrity constraint. This is quite easy, as it just the ordinary integrity constraint for simple past events:

The query ?*Happens(see$_e$,t)*, $t < now$ `succeeds`, and the query ?*HoldsAt(f$_1$,R)*,

---

[105]e_{start} is any event that causes readiness to begin. Supposes Mary is crossing the street and I turn my head and see it.

[106]For atelic events, any arbitrary *stop* of the complement will do.

[107]Note that as the complement event is in simple aspect, it is also defined as a hierarchically planned event that does include a $finish$ event that stops the activity fluent and triggers the result state of having finished the accomplishment.

[108]From the speakers point of view, a *stop* event has happened, which may in reality only be a brief pause, or just enough *running* for the speaker to view it as a *running* event.

. . . , *HoldsAt(f$_n$,R), R$\geq$ now, Happens(see$_e$,R)* `fails`.

As *see* is defined to be a hierarchically planned event, it not only takes time, but the *seeing* is bounded by the time of the complement event, i.e. there is no seeing after the complement event. Once the complement event finishes, there is nothing to drive the *trajectory* clause of *see*, and so the query ?*Happens(see$_e$,t)*, *t < now* fails as the hierarchically planned definition of *see* requires that the fluent *see* holds at *t*, which can only be the case when the complement event is happening.

Past tense *see* with a progressive complement

This is the case with sentences such as *I saw Mary crossing the street* or *I saw the earth spinning on its axis*, which have no implications on the complement except that the associated activity occurred. This tense/aspect combination is virtually the same as that for present tense *see* with a progressive complement, except that the integrity constraint is located in the past instead of the present. There is one possible difference. I am unsure as to the felicity of continuing *see* into the future. For a contrived example, I could be on a surveillance assignment and asked what I saw Mary doing a minute ago.

(57)   When I first looked through my binoculars, I saw Mary standing by the window, and I still do.

More likely I would say, 'she still is standing there', as my seeing it is implied. Nevertheless, this type of formulation is possible, whereas if I say, "I saw Mary open the door a minute ago", I cannot still be seeing it. Now, if something like sentence (57) is acceptable, nothing more needs to be done to the formalisation. But, if seeing is restricted to the past, even for a progressive activity, then we can copy most of the scenario for the present tense version, but treat *see* as a simple past tense event. That is, we turn it into a hierarchically planned event with the corresponding integrity constraint. The former is preferable to me on aesthetic reasons. The other cases would have to be created as sort of parallel reflection of the complement event in the event structure of *see*. It would be nice to keep this parallelism. For the case where *see* can extend arbitrarily in the present, the formalisation[109] is exactly the same as for present tense *see* with a progressive complement, but for the location of *R*. In this case, the integrity constraint is:

?*HoldsAt(see,R), R <now* `succeeds`

---

[109]And associated scenario for the dynamics of *see*.

Been caught borrowing

The final example for *see* is a sketch of how a derivation of a 'mistaken' activity would work. Recall the variant of an example from (Gee, 1977)

(58)   I saw John stealing the car, but at the time I thought it was only bor-rowing.

This is said from the perspective of the speaker recalling the (complement) event and changing his mind about what it actually was.[110] For this context, there is no spatio-temporal difference between *steal* and *borrow*. Suppose John is the brother of the car's owner, and took it without his permission. A third party observer who knew this much would naturally see it *as* borrowing at the time, only later finding out that John and his brother had a blazing row and this was John's revenge. It is this sort of knowledge update that changes the appropriate characterisation of the event.

   For this context, we will assume that the relevant difference between a *stealing* and *borrowing* event is the fluent *permission*. If it holds at the time, the event is *borrowing*, if it does not, the event is *stealing*. Furthermore, as we are speaking of brothers, we can give a constraint on permission, such that if something (like taking the car) is not forbidden, then there is permission.

$$\neg HoldsAt(forbidden, t) \rightarrow HoldsAt(permission, t)$$

   Importantly, if the speaker has no information one way or the other, then by negation as failure, we get $\neg HoldsAt(forbidden, t)$, meaning that a query beginning with $?HoldsAt(permission, t)$ will then succeed. Now, the two events can be defined in terms of these fluents plus a spatio-temporal bit they have in common.

(59)   a.

$$Happens(take, t) \wedge HoldsAt(forbidden, t) \rightarrow Happens(steal, t).$$

   b.

$$Happens(take, t) \wedge HoldsAt(permission, t) \rightarrow Happens(borrow, t).$$

At the time of the theft (where the speaker sees only borrowing), the query for the complement event will be

---

[110]At the time he would have said, 'I see John borrowing the car.'

$?Happens(e,t),\ t < now$

with parallel unifications for $e = borrow$ and $e = steal$.

Which event gets substituted depends on the status of $?HoldsAt(forbidden, t)$. If this query succeeds, the substitution $e = steal$ succeeds, and $e = borrow$ fails. If $?HoldsAt(forbidden, t)$ fails, $e = borrow$ succeeds and $e = steal$ fails. Thus, in a minimal model, with no information about *forbidden*, the complement resolves to borrowing, but in an extension the complement may resolve to stealing. This is exactly what happens in sentence (58), where the speaker is looking back on the event with an update of knowledge, in this case that John was forbidden to take the car.

4.5.3   Watch

As seen in section 4.2, *watch* is not nearly such an event-type chameleon as *see*, but is rather an unabashed activity. It behaves like a typical activity in requiring the progressive for non-narrative present tense, and is incompatible with uncoerced achievements or points. A major difference between the two verbs is also the interpretation of their complements. Where the 'progressive hypothesis' certainly holds for *see*, the semantic aspect of the complement is sometimes dependent on the aspect of the matrix verb. The difference can be seen in the following:

(60)   a. I saw Mary cross the street.
        b. I watched Mary cross the street.
        c. I was watching Mary cross the street.

In the first two sentences of (60), the simple aspect complement is indeed perfective and completive, with the strict implication that Mary made it to the other side of the street. In contrast, in sentence (60-c), thought the complement is also simple aspect, the implication is only that Mary was crossing the street, with no implication of completion.[111] It seems then, that though there is a simple aspect complement in (60-c), it nevertheless has an imperfective interpretation. This phenomenon is also most likely responsible for the observation that in both

(61)   a. I am watching the light flash.
        b. I watched the light flash.

---

[111]There is also the less stylistically preferred *I was watching Mary crossing the street*, which has the same meaning, but some speakers find the double-*Ing* construction rather awkward.

the light must flash more than once.[112] Informally, it is best to view *watch* as
shining a spotlight on an activity for at least a small amount of time, making an
analogue of seeing something in an instant impossible.

The second major difference between the two verbs of perception is that there
is a more 'intentional' nature to *watch*. This can be seen by comparing the
following two questions:

(62)   a. Why did you see the young children playing?
          b. Why were you watching the young children play?

(62-a) is a rather odd question, and would normally be answered as if it were a
'how' question, i.e. I was walking by the playground and looked up, etc. In a
contemporary climate the question in (62-b) has a more sinister implication, and
had better be answered in a way that one is innocent in intention. Of course, one
could be forced (in a Clockwork Orange sort of way) to watch something,[113] so
intention is not quite the correct formulation. Rather, the this difference between
see and *watch* is a matter of paying attention in the latter case.

Finally, there is the implicational chain for *watch* that needs to be taken into
account. A sentence like

(63)   I am watching Mary cross the street.

implies both that I see Mary crossing the street, and (assuming a default veridi-
cality) that Mary is crossing the street. Thus it can be concluded that watching
something is seeing it *and* paying attention to it for a given amount of time.
Indeed it is just these insights that are needed for the formalisation, along with
taking the imperfective aspectual shift into account.

The basic scenario

While *watch* does share much in common with *see*, the previous discussion makes
clear the major differences. The major distinction is *watch*'s activity aktion-
sart, coupled with the notion mentioned in the last section of *watch* shining a
metaphorical spotlight on an event. Thus the scenario of *watch* will differ from
that of *see* in two major ways. The first is that the verb itself, is always a fluent,[114]
and moreover not a parameterised state fluent, but one denoting an activity.

---

[112]This aspectual shift is not confined to *watch*, but may be something more general that
occurs with a progressive matrix verb and simple aspect complement. In Chapter 5, it is
observed that one of the readings of *I'm making the puppet walk across the stage* has the same
sort of imperfective interpretation of the complement.

[113]We can also be internally compelled to watch something horrible ('I just couldn't turn
away'). Whether this is intention or not is a matter I shall leave to the psychologists.

[114]As opposed to possibly being a punctual event, which is a special case of *see*, e.g. *I saw a
flash.*

Thus the constraint on *watch* that will serve as an analogue to the stative *ready*, is an activity fluent *attn*. That is, one cannot merely happen to watch something (for a brief instant) by accident, but rather must put some effort into the process. Certainly, one talks of *paying attention*, i.e. doing something, and in the case of *watch*, this could mean having one's eyes follow a scene, or simply staring intently. Of course, such activities do take some time, thereby ruling out the possibility of paying attention to an instantaneous point.

Thus, the first ingredient needed is a fluent *attn* which is the formal analogue of the attentional nature of *watch* described above. Secondly we shall use the version of *see* defined in Section 5.2 that models such sentences as *I see Mary crossing the street* or the perfective analogue, *I saw Mary cross the street.* We shall refer to this as see_f, i.e. seeing an activity, with $watch_f$ is defined as an f_1 activity fluent as follows:

$$HoldsAt(attn,t) \wedge HoldsAt(see_f,t) \rightarrow HoldsAt(watch_f,t)$$

with the constraint on the fluent *attn* being similar to that of *ready*:

$$?Terminates(e,attn,t), \neg Terminates(e,watch,t) \texttt{ fails}$$

meaning that no event can terminate one paying attention (to something) without also terminating watching it.

As the definition of *watch* is dependent upon *see*, the scenario of the complement of *watch* is already taken care of. The completion of the activity denoted by the complement (unlike with simple aspect *see*) depends both upon the aspect of *watch* as well as on the aspect of its complement. When *watch* is in the progressive aspect, these completion conditions are not an issue an we shall thus start with the simpler construction first.

Watching

As *watch* is a simple activity, the standard condition that *watch* must appear in the progressive in non-narrative or habitual present holds. Moreover, as detailed in 4.5.3, the morphological aspect of the complement is irrelevant. Whether morphologically simple or progressive aspect, the complement activity is always interpreted as 'in progress' at the reference time. For a concrete example, we shall return to

(64)   I'm watching Mary cross the street.

Using the fluent *cross* which denotes the activity of a *cross the street* accomplishment, *watch* can be defined as follows:

$$HoldsAt(attn,t) \land HoldsAt(see_{cross},t) \rightarrow HoldsAt(watch_{cross},t)$$

The only thing remaining is then to give an Integrity Constraint for the present progressive in (64).

$?HoldsAt(watch_{cross},R)$, $R{=}now$ `succeeds`

This Integrity Constraint begins a derivation with the conditional above, causing the database to add to the discourse model both that the activity portion of the complement[115] is seen (and is also imperfective) and that attention is being paid to it. A property of this construction is that *watch* and *see* may be extensionally equivalent, but are intentionally distinct due to the addition of the *attn* condition.[116]

Finally, there is the manner of implication to the perfective to examine. While sentence (64) has no strict implications for either the complement activity completing or watching it until completion, it is the case that if all goes well, I will have watched Mary cross the street. The telicity of the latter is taken care of in by the occurrence of see_f, which calls upon the complement's scenario in its definition. What remains then is to define what it is to watch something until the end.

Watched

There are two cases to consider here, exemplified by the following two examples:

(65)   a. I watched Mary cross the street.
         b. I watched Mary crossing the street.

In both cases, the watching is confined to the past, with the difference lying in the status of the complement event. In sentence (65-a), Mary successfully crosses the street, while in the second case, the status of her safety is left open. Should the complement be an atelic activity rather than an accomplishment, the only difference would be the progressive rendering could continue into the future, while the simple aspect version will have stopped.

Beginning with the case in (65-a), there are two event completions to be accounted for. The first is that the complement event does finish, and the second is that it is watched until the end.[117] So, like the treatment of past-tense *see*, there will not be a partial object counterpart to *watch* in a standard accomplishment

---

[115]This is when the complement is telic, as in *cross the street*. Should the sentence be *I'm watching Mary run*, there is only the activity to worry about.

[116]The extensional equivalence may be partial. An interval of *watch* will always be an interval of *see*, but the converse may not be the case. I can see something happening, and find it interesting enough to then begin watching it.

[117]Though as with *see*, it needn't be watched for the entirety of the complement event.

dynamics, but rather a construction that connects the canonical events of watching to the complement event, which is where the true accomplishment dynamics lie. That is, we need to define *start* and *stop* events for *watch* that will allow us to build a perfective, hierarchically-planned event. This is seen in the following.

$start_{watch}$ is defined in two separate clauses:

- $HoldsAt(see_f,t) \land Happens(e_{attn},t) \rightarrow Happens(start_{watch},t)$

- $Happens(start_{see},t) \land HoldsAt(attn,t) \rightarrow Happens(start_{watch},t)$

The first clause handles the case where one sees something already happening and it captures there attention enough to begin watching it – the event e_{attn} is merely any event that begins the *attn* fluent. The second clause is a rather special case. Imagine a spectator avidly awaiting the beginning of a race and already paying attention to the runners in the starting blocks. As they are still motionless, one can neither see, never mind watch them running. Once the starting gun goes off and the contestants start running, this is then seen, and since *attn* also holds, watching also begins.

The finish event is a construction very similar to that for the complement of *see*. Finishing watching something requires both that the complement event completes,[118] and that it is watched up until the end. As with the finish event for perfective versions of *see*, this is defined by two clauses. Taking $finish_c$ to be the telic culmination point of an accomplishment, the two clauses are

$Happens(finish_{watch},t) \rightarrow Happens(finish_c,t)$
$Happens(finish_{watch},t) \rightarrow HoldsAt(attn,t)$

Putting the ingredients together, we can then define a hierarchically planned event for a sentence like (65-a).

$Happens(start_{watch},s) \land Happens(finish_{watch},r) \land s < t \leq r \land HoldsAt(watch_f,t)$
$\rightarrow Happens(watch_e,t)$[119]

The integrity constraint for past tense is of the standard form

---

[118]If the event is not an accomplishment, but an atelic activity it is enough that it merely stops at an arbitrary point.

[119]Should the complement event be atelic, the finish event associated with the complement can be replaced by an arbitrary stop event.

The query $?Happens(watch_e,\text{t})$, $t < now$ `succeeds`, and the query $?HoldsAt(f_1,\text{R})$, $\ldots$, $HoldsAt(f_n,\text{R})$, $R \geq now$, $Happens(watch_e,\text{R})$ `fails`.

The case of (65-b) confines the watching, perfectively to the past, but says nothing about the activity either completing or continuing on into the future. Thus, the only modification that needs to be made is for the end of the event. For this an event, $stop_{watch}$,[120] is needed which is merely an event that terminates $watch_f$, i.e. an event that terminates either $attn$ or see_f and says nothing about the complement event completing or stopping. This hierarchically planned event is then:

$$Happens(start_{watch},\text{s}) \wedge Happens(stop_{watch},\text{r}) \wedge \text{s} < \text{t} \leq \text{r} \wedge HoldsAt(watch_f,\text{t}) \rightarrow Happens(watch_e,\text{t})$$

whose tense is anchored also by the integrity constraint above.

4.5.4 Comments on the formalisation and syntax

Despite the criticism of Felser's analysis in the preceding sections, there is only one major point of difference, and one minor. Certainly, there are generally syntactic elements of restrictions on the complement (i.e. no complementizers, particle movement, etc.) that are captured perfectly. Moreover, that she treats the complement as a constituent, rather than an argument structure mismatch (as discussed in Section 3.1)[121] simplifies matters considerably. We have also taken the notion of linking between the two events rather seriously. However, this is not done to merely guarantee temporal simultaneity, but to derive some of the more subtle aktionsart phenomena discussed in Section 4, where the temporal profile of the perception event is partially defined and dependent upon the nature of the complement.[122]

The two differences are of semantic interpretation and that of how tense is handled. The first has to do with her incorrectly ruling out perception complements with 'individual-level' predicates. This has nothing to do directly with her syntactic analysis, but rather her not allowing *all* verbs to possibly have stage level interpretations. As seen in this chapter, as well as Chapters 2 and 3, this is empirically false. But,[123] should such an interpretation be allowed, verbs such as *know* would then be allowed to have a syntactic representation that would make them felicitous in direct perception reports.

---

[120]No pun intended.

[121]As well as reviving Fillmore's progressive hypothesis.

[122]The sort of 'event control' goes both ways in the EC analysis, which I think is justified when the aspectual properties of reports with *watch* are taken into account.

[123]Still speaking in terms of event variables and positions for the sake of argument.

The second issue is more difficult, and is discussed in further detail below. It is simply that *ing* forms may indeed carry their own tense in certain constructions, which does cast doubt on the syntactic analysis.[124] I have no answers to this problem, but will offer the barest of comments and suggestions in the next section. Originally the plan was to use the syntactic analysis as a platform to explore solely the interactions between event-type, tense and aspect. From the syntax, the complement as an independent event, whose tense overlapped with the complement came for free from the syntactic analysis. At present, tense is anchored semantically,[125] which conceptually doesn't seem particularly out of order. Nevertheless, the original plan was to only use an alternative event semantics that is particularly suited for modelling the phenomena seen in Section 4. As a preview of the tense conundrum, here are two (simplified) structures (Felser, 1999, p. 119)[126] that show the different tense relations between a simple Asp and a futurate infinitive.

(66)   a. John $[_{T'}$ is $[_{AspP}$ $[_{Asp'}$ looking$_i$ $[_{vP}$ the reference...t_i up...
         b. We expect John $[_{T'}$ to $[_{AspP}[_{Asp'}look_i$ $[_{vP}$ the reference...t_i up...

The first is an example of the present progressive,[127] where the time is only present time. For (66-b), *expect* is indexed for the present tense, but since the infinitival has its own tense head, the complement tense is not controlled by the matrix, but rather by the infinitive, and has a future interpretation.

This patterning works fine with *expect*, but with *intend* it breaks down for many English dialects. This matter and others are discussed in the final section of the chapter.

4.6   Coda – Prepositions vs. *ing*

In section 3, a bit of doubt was cast on the universality of the simultaneity principle. Certainly, for direct perception sentences the is a temporal linkage between the matrix verb and the complement. For Felser, this was due to the complement being an aspectual phrase and having no tense of its own. We saw that *see* itself is no guarantee of this, as one can use it in a non-direct perception

---

[124]Not only hers, however. In the next section we will see that this also contradicts some authors in the functional tradition as well.

[125]This is done with clauses that include both the complement event and the seeing event with identical timepoints.

[126]Example (66)-(b) is modified in that the matrix verb is put in the present tense.

[127]How a futurate progressive of *John is looking the reference up* will work under this analysis is a mystery to me as the time of *is* and the time of the event will differ, unless a preparatory event is accounted for somewhere in the semantics. I think the solution to the problems pointed out in more detail in the next section, will partially involve teasing out the difference between the *real* present progressive interpretation and a future oriented one.

predictive sense that places the complement event in the future. She extends this analysis to aspectual verbs as well, (p. 190)

(67)   a. Sue continued filling in the form.
        b. Bill started fixing his car.
        c. Mary stopped cleaning her room.

where the complements of the aspectual verb are aspectual phrases, and indeed, the complement is simultaneous with the aspectual event (e.g. *stop*, *continue*, etc.) For some reason, she neglects to include the *ing* phrase having an overt subject, such as

(68)   Bill stopped Mary cleaning her room.[128]

The difference is important, however. The first type are 'raising' predicates, where the subject of the matrix verb is co-identified with the subject of the complement. This alone, doesn't guarantee simultaneity as *intend* is also a raising predicate,[129] but in non-American UK derived dialects there is no necessary simultaneity between matrix and complement.

(69)   a. A priest who regularly conducts Mass for Tony Blair and his family refuelled speculation last night that *the British Prime Minister intends becoming a Roman Catholic.*[130]
        b. He intends *taking legal action.*[131]

Matters become more complicated when we look at non-raising examples of aspectual verbs, such as in sentence (69). It is dialect dependent on whether they obey the simultaneity principle. This is in contrast also to a functionalist syntax point of view, as is illustrated in (Valin & La Polla, 1997, p. 472).

(70)   a. Robin stopped Kim singing 'Advance Australia Fair'.
        b. Robin stopped Kim from singing 'Advance Australia Fair'.

Example (70-a)[132] is *marginal* (emph. added) for many speakers, but nevertheless the contrast relevant to this discussion comes through clearly. In (a), Kim is singing the Australian national anthem and

---

[128]It could be that she is focussing on a standard American English dialect, where such constructions are less felicitous than more UK derived dialects.

[129]Though not an aspectual verb, of course.

[130]Irish Independent, 15 October 2004.

[131]As noted previously, in American dialects it is preferable to use a *to* infinitive or a preposition such as *on* – i.e. *She intends to take/on taking legal action.* This example comes from a BBC radio news report.

[132]This is my numbering obviously.

Robin stops her; since the singing has been going on at the time of the stopping action, the two overlap temporally. In (b), on the other hand, no singing ever occurred; that is Robin acted to stop Kim before she even started to sing the song.

He then gives a more generalised formulation for the linking between matrix and complement

a. Constructions with zero marker on linked unit: [+temporal overlap]
b. Constructions with *to/from* marker on linked unit [-temporal overlap]

Indeed, this does work when the aspectual verb is of the 'raising' kind; the difference can be seen via the contrasting interpretations in infinitives and *ing* complements.

(71)   a. Mary started to drive to work (but before she got in the car she changed her mind and took the bus).
        b. Mary started driving to work (*but before she got in the car she changed her mind and took the bus.)

But once there is an intervening NP between the aspectual verb and the *ing* complement,[133] simultaneity may merely be an accident of dialect. The following is taken from an article in the Observer.[134]

For nearly 20 years, Sin Phillips was in thrall to Peter O'Toole, the man she married and describes now as 'a dangerous, disruptive human being'. She loved him to distraction, but, whether deliberately or not, *he stopped her doing the one thing she had always wanted to do.*
'Of course I could have been more successful. I would not have lost that crucial decade where you should be laying down a body of work... *I never stopped working, but I wasn't doing what I wanted, serious work at the RSC.* I was extremely occupied with being married and being a mother and fitting in with a very, very successful career.'

As can be seen from the above, it is not the case that she was ever doing what she wanted to do (work at the Royal Shakespeare Company). For a group of British speakers, at least, (70-a) can also be interpreted as (70-b). I suspect

---

[133]Something that is necessary for verbs of perception with aspectual complements. *I saw Mary crossing the street*, but not **I saw crossing the street*.
[134]Sunday July 29, 2001.

this is an *-ing* form used as a futurate or at least for what is planned or intended. Simultaneity can then be saved (in a sense), but only with a broader view of the event.

This is no anomalous example. The following is from the City of Edinburgh web-page regarding the Victorian body snatching controversy.

(72)   Mrs. Gray returned in the morning to find Burke acting very suspiciously, and he *stopped her going over to the bed* to collect her child's stockings.[135]

There is also the verb *remember*, which is examined in (Higginbotham, 2003). He calls complements for sentences like *He remembers giving the speech*, gerundive,[136] but also that they denote events. The paper is partially an excuse to explore the contribution of PRO. But there are differences between the PRO version above, an aspectual phrase and a real gerund. This is Higginbotham's pair.

(73)   a. I remember saying that John should finish his thesis by July.
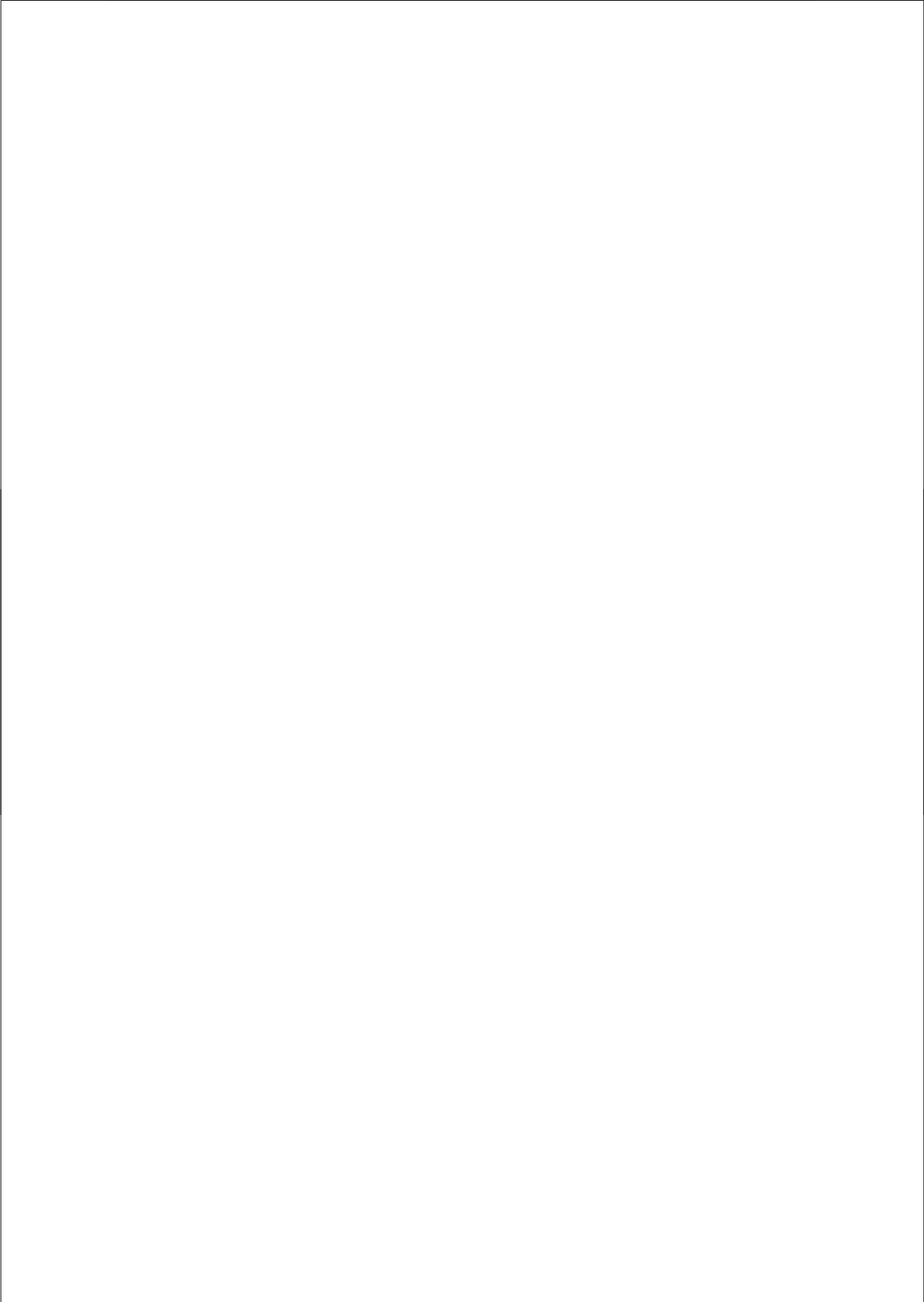   b. I remember my saying that John should finish his thesis by July.

He rightly notes that in (73-a) it is not only the event that is remembered but it is remembered from the point of view of the speaker being the agent of it. This isn't quite the case with (73-b). In any case, semantically, there can be no simultaneity of course. This is less of a problem for Felser, as they are gerunds and not aspectual phrases. But the point was brought up to show the difference in other cases, between the use of PRO and an explicit subject.

While I have no answers for any of this, my guess is that a variety of factors involving both argument structure and lexical semantics[137] are at work, along with optionality for different dialects. Chapter 6 will explore in greater depth the profound mystery of verbal *ing*. This too may brings up many new questions.

---

[135]As there is no additional context as in the other example, the situation is no entirely clear. Mrs. Gray could have been going over to the bed and stopped her (temporal overlap) or could have ushered her out of the room before she had a chance to begin going over to the bed (no temporal overlap). In either case, this was probably her intention when she entered the room.

[136]Strictly speaking, they cannot be aspectual phrases as they allow a possessive as the 'subject' of the *ing* complement.

[137]When an aspectual verb is used in its raising construction, simultaneity happens, but not in the case of *intend*, which is, semantically quite a different animal.

Chapter 5

Causatives

5.1 Introduction

5.1.1 What are causatives? The general notion

In a general way, the notion of causation is crucial for the representation of semantic structure. Ordinary transitive sentences that denote accomplishments or activities, such as *John built a house* or *Bill threw the ball*, while not considered 'causatives' undeniably have a notion of causation at their core. The first sentence represents a verb of creation, where John causes something to come into existence, while in the second, there is a force transmission of Bill to the ball that causes the ball to move.

(Croft, 2000b), revising his notion of causal chain from (Croft, 1991), stresses the importance of capturing force-dynamic and causal relations between events along with aspectual contours of the events. This especially crucial for argument linking. However this notion had presented him with a number of problems.[1] His solution is rather elegant, and can be illustrated by the quote that follows.

> ... these problems can be solved in a single, disarmingly simple stroke. A complex verbal semantic structure is decomposed into subevents such that each subevent involves only one participant and its own aspectual contour; and the subevents are related to each other causally. (Croft, 2000b, Chapter 3, p. 6)

Serendipitously, this is the basic strategy of the Event Calculus as well, and while in the formalism later in the chapter we have not formulated an explicit method of argument linking, we do use the spirit of the above quote as our guide, as the notion does form the heart and soul of the Event Calculus, as far as aspectual and aktikonsarten representation is concerned.

---

[1] See (Croft, 2000b) for a list of them.

While causation can be said to be a core facet of many verbal profiles and event types, there are also, in the literature, specific constructions that are appropriately named causative constructions. For (Song, 2001), the following sentence can be seen as a prototypical example of a causative construction,

(1)   Elizabeth made the chef eat the leftovers.

where Elizabeth did something that has a result that the chef carries out an event, in this case eating the leftovers. Importantly for Song, the cause is expressed by an abstract predicate (*made*), while the activity of the chef (*eat*) has far more specific lexical content. He notes (citing (Kemmer & Verhagen, 1994)) that all that is expressed by verb representing the cause is 'the pure notion of cause...without more specific lexical content.' (Kemmer & Verhagen, 1994, p. 117)

Moreover, the causer NP and the causal predicate are 'foregrounded' in respect to the causee and his event. The causer is the subject NP of the construction, and the predicate is inflected. Contrastly, the predicate of the caused event is non-finite, or what Song refers to as 'not a full-fledged verb' (p. 258). He contrasts (1) with the following, that, while certainly causal in nature, is not an instance of the causative construction.

(2)   The chef ate the leftovers because Elizabeth burnt his meal.

Here two substantive lexical verbs are used to indicate the separate events (and are grammatically independent of each other), and the information regarding the causer is placed in the background (in the 'because' clause) rather than fore-grounded as in (1). Finally, the cause in sentence (2) has specific lexical content rather than the abstract causal predicate in sentence (1).

However, it is not the case[2] that causative constructions can only use 'abstract' predicates such as *make* or *let*. While *burn* certainly has lexical content, the following sentence would also be considered a causative construction.

(3)   Elizabeth burnt the meal.

A crucial distinction between sentences (2) and (3) is the relationship between the causal event and recipient of the causality. In (2), the causal recipient is the chef and we are given a great deal of lexical specification of what the casual event was. In sentence (3), the causal recipient is the meal and while *burn* certainly has lexical content, this content indicates only what happened to the recipient of causation. Whatever act Elizabeth does (or fails to do) to cause the meal to burn is completely abstract. While we may stereotypically associate certain actions that Elizabeth could have done in (3), the causing event can be nearly anything, and in this sense, the causing event in (3) is just as abstract as if *make* were used. With this in mind, we shall be focussing on causative constructions

---

[2]Song does certainly look at lexical causation.

exemplified by the first and third sentences, and ignore the more general 'causal' constructions such as sentence (2).

5.1.2 Causation

The notion of cause is an incredibly rich and complex notion and can be looked at from scientific, philosophical or psychological perspectives, to name just a few. Instead, we shall follow (Kemmer & Verhagen, 1994) in restricting the notion of causation to its manifestation in human conceptualisation.[3] What is important is how the speaker views a given situation. Whether or not a speaker views a given situation as causal, and in what way it is causal will go a long way in determining what kind of linguistic structure they will use to express their view of the situation. Moreover, allowing for construal differences in the area of causation and the participants involved allows us to incorporate into a semantic theory many creative extensions of certain constructions, violations of defaults, and changes in the language over time in regard to causative verbs and constructions.

5.1.3 Plan of chapter

We begin by looking in the next section at general conceptual notions of causation and their semantic typology. Section 2.3 briefly brings us to a basic look at English causatives followed by sections 2.4 and 2.5 which look to cross-linguistic typology to broaden the notion of causation types and understand the various possibilities that are possible in conceptions of causation. This allows us to discover types, which while not explicitly obvious in English (on the surface) do indeed manifest themselves. After section 2.5, which brings the previous two sections together, we enter into the preliminaries of the formalisation of some types of causative situations for English. This entails a close examination of the role of the participants and its effect on the implications for tense and aspect constructions, followed by the actual formalisations. The last two sections are rather short and serve as codas to the previous sections. Section 3 is a quick look at the change in the applicability of lexical causatives in English diachronically, and Section 4 is an all to short look at resultative constructions, especially when they involve lexical causatives. While short, it provides a good look at the importance of taking construal rather seriously, as we shall see some rather creative (and technically 'ungrammatical') uses by speakers that only make sense if one allows for the speaker's view along with the general causal meaning of many types of resultatives[4] to interact felicitously.

---

[3]They expand on this, saying that 'the human conceptualisation of causation...must be based in some fundamental mode of modes of chunking and organising perceived reality that allows humans to interact successfully with their physical and social environment. (p. 117)

[4]As well as the 'X's way' construction.

5.2   Causatives

This section explores causatives from a variety of angles. We shall first look at
what causation is considered on the semantic level, followed by a brief survey of
English causatives and how they exemplify the causative concepts. The reader will
then be asked to leave familiar territory and embark on a typological exploration,
which is useful in revealing subtleties about causatives that are not immediately
apparent from the English examples.

These subtleties will then be shown to also exist in English, not through
morpho-syntactic marking, but rather through contrasting implications depend-
ing on the various causal scenes. With these distinctions in mind, only then can
a proper way of formally describing causative structures in English be made.

5.2.1   Semantic Parameters

(Kemmer & Verhagen, 1994) give a series of sentences that are illustrative of the
various semantic intricacies involved in causation and are repeated here for the
benefit of the reader.

(4)   a. She made it fall over by pushing it.
         b. She made it fall over by rolling the ball into it.
         c. She made him cough.
         d. She made him type the letter.
         e. She had him type the letter.
         f. She let the water run out of the bathtub.
         g. She let him eat some brownies.

These are typical examples of periphrastic causatives. It may be useful for the
reader to first imagine what the different causative possibilities are, depending
both on the auxilliary verb used as well as the situation denoted by the comple-
ment. Imagine, especially, the different possible ways that the cause produces the
effect, temporal linking, and control relations between the participants.

Causation Types

(Song, 2001) provides a description of various causation types that are used to
then identify and differentiate different morpho-syntactic causative constructions
cross-linguistically. In essence, a transitive or ditransitive causative[5] expresses a
complex situation with a causing event, initiated by a causer in order to bring
about the caused event. In the caused event, there is a causee who either carries
out an action or undergoes a change of state. Observe the basic examples

---

[5]With intransitives there is one event ignored.

(5)   a. Mary broke the vase.
         b. John made Bill turn on the light.

In (5-a) Mary is the causer who either does (or can be held responsible for) a certain event that causes the vase to undergo a change from whole to broken – dropping it on the floor, for example. In (5-b), John is the causer who manages to have the causee (Bill) perform the action of turning on the light.

This brief description seems simple enough but hides two distinct levels of description – causation between the events and causation between the participants. First, there is a distinction to be made between whether the relationship between the two events can be seen as direct or indirect causation. That is, can the event associated with the causer be directly associated with the causee event? Taking the example in (5-a), a typical scenario would have Mary knocking it over thus causing it to break. This would typically be considered an example of direct causation, as the two events are temporally linked and there is no intervening event in the vase breaking.[6] In other words, direct causation can be seen as a causative situation where both events temporally abut with no intervening events playing a role.

Indirect causation, as expected, is the opposite of the above. Imagine (5-b) as describing a situation where John put an obstacle in front of the doorway knowing that Bill always rushes into rooms regardless of the light situation. Bill falls and realises that he should turn on the light. Here John is nowhere about and his sole event contribution (placing the obstacle) could have occurred at any time before the caused event. Obviously then, the causing event and caused event are separated by time, requiring the intermediate events of Bill entering the room and falling down to carry out the caused event. However, in order to see John as the causer (and hence as seen to have done something responsible for the caused event) the caused event must be temporally connected back to the causing event though a chain of events. In future I will refer to the level of causation as 'granularity', which shall be explored informally in a following subsection and then return to play a role in the formalisation. Moreover, cross-linguistically there are markedness differences depending on whether the cause is direct or indirect.

The reader should note that (4-a) is the only unequivocal example of direct causation, while (b) and (e)-(f) are all examples of indirect causation. (4-c) and (d) can be interpreted either way, but the more natural way is to assume that they are more compatible with a direct causative reading. Note that under a direct causative reading of (b), it is the case that another distinction is necessary. Making something fall over by pushing it and making someone cough by telling them to are both direct causation, but the way the participant brings the cause about hints at the need for a deeper analysis.

---

[6]This is not a claim that (5-a) always involves direct causation, but only that the typical scenario given is such an example.

The second distinction, that of the level of description between the causer and causee often has an effect on case-marking or choice of causal verb. The distinction to be kept in mind here is what is called either *manipulative* or *directive* causation. Typically, manipulative causation involves the causer acting physically upon the cause, while directive causation relies on other, more subtle, means.

Song gives a nice example of John causing Bill to fall by rolling a shopping trolley into him. Given the distinctions made in the last paragraph this is an example of manipulative causation. But unlike our scenario for (5-a), it is also an example of indirect causation. John acts physically to affect Bill, but there is an intervening event, namely that of the shopping trolley rolling into Bill that is needed to make the causal chain complete.[7] Thus there is the possibility to have manipulative causal situations that are either direct or indirect, reinforcing the need to analyse the situation on two distinct causal levels.

Directive causation is associated with non-physical means that the causer may use in order to affect the causee. Song's example involves a doctor telling a patient to lie down. Here the causal event is non-physical, but may be seen nevertheless as an example of direct causation. But if the doctor told his nurse to tell the patient to lie down, it would be an example of both indirect and directive causation. As expected, the difference between the two types of directive causation becomes manifest when we look at the ways the two types would generally be expressed.

While ostensibly a topic in another area, the cross-linguistic section will show that the worlds languages take account of the levels of causality more explicitly than shown so far. The type of causative situation will often determine the way it can be expressed, with direct types of causation usually corresponding to less marked or direct linguistic forms. We have already observed types of situation that are both direct and directive or indirect and manipulative. The various combinations along the two dimensions allow for a good deal of variation in languages that have the morphosyntactic ability to do so. Languages such as English (or Dutch) may express the differences lexically,[8] but often this morpho-syntactic poverty can lead to an initial confusion when analysing causality. For example, the ambiguity of 'She made him cough' revels in the various ways the two dimensions of causality can combine.

Kemmer and Verhagen (1994) further elaborate on the participant causal distinction by noting that both *enablement* and *permission* (the former physical and the latter social or inaction) can be also seen as causation types, where it is even

---

[7]One might argue that Mary dropping the vase, thus breaking it, also involves an intermediate event, namely the vase falling. A crucial difference is that this purported intermediate event is part of what happens to the vase, rather than intervention by a third object. The difference is one of perspective, but it crucial for understanding how causality is construed. Should the reader still be sceptical, imagine that Mary breaks the vase by smashing it, thus disallowing any third party mediation.

[8](Kemmer & Verhagen, 1994)) note that *laten* is used more for indirect (and often directive) causation, while *doen* is more direct in nature.

possible the causer may have no intention of having the caused event come about. For example, if I see someone walking near a precarious cliff and do nothing to warn him, my inaction can be said to be a cause of him falling off. Similarly if I have removed the fence thus causing him to fall off, it can be seen as a type of *enablement* causation.

Now, despite the possibility of ambiguity conceded above, the distinction just made between more intentional forms of manipulative and directive causation does sometimes manifest itself the very non-exotic language of English. Going back to Kemmer and Verhagen's examples in (4), those with 'make' are much more likely to involve manipulative causation (telling someone to cough, or hitting them on the back, for example), while 'having' someone type a letter is more likely to involve the intermediacy of someone else or the willing participation of the causee.

Before moving on, the reader may have come up with a variety of situations to fit the first example of the section and then, upon reflection, wondered whether a certain situation was direct or indirect, manipulative or directive, and perhaps even pondered over whether it matters if the causee had any autonomy in performing the caused event. For example, a bare situation such as causing someone to type a letter can be construed in various ways. This, and if the causation is construed as direct or indirect affects the linguistic form used. That is, the same real world situation can be seen in various ways, the way in which a speaker desires to present the situation affects the way the utterance will be coded. Much more will be said of this in upcoming sections, but first we shall examine in detail two semantic notions that will be highly relevant to the formalisation.

Notion of external event

From the above, it is clear that the idea of causation can be realised in a number of ways. One of the basic notions is that where a causative is used with a causer argument (e.g."Sunshine" in "Sunshine melted the ice" or a periphrastic version using 'make') there is an external event that causes the caused event. In this caused event the participant can either be the agent of the caused event (e.g. ordinary uses of "John walked the dog", or undergoing a process, as with "melt".)

On the verbal level, (Levin & Rappaport-Hovav, 1995) make a distinction between internally and externally caused eventualities. An intransitive verb that describes an internal eventuality can be said to be describing "some property inherent to the argument of the verb" that is "responsible" for bringing about the eventuality. With agentive verbs ("play" or "speak"), the internal cause is the will or intention of the agent. But, they note lack of control by an agent does not mean that the eventuality is not internal. They cite the verbs "blush" and "tremble", which while not agentive, can be said to arise from the internal property of the subject (in these cases an emotional reaction).

Externally caused verbs (as the name suggest), are those verbs whose eventualities are caused by something possibly different from the object that has some

degree of control in bringing about the eventuality. *Break* is a typical example of such a verb. While it is natural to describe something either 'breaking itself' or 'breaking by itself', this is a case of the object being both the external cause and the causee. In fact, when a verb like *break* is used intransitively, "our knowledge of the world tells us that the eventuality these verbs describe could not have happened without an external cause." (p.93).

The point of this is that (Levin & Rappaport-Hovav, 1995) proposes that the difference between internally and externally caused verbs is that only the latter can participate in the so-called transitive alternation. Internally caused verbs can only appear intransitively or can be seen to be caused in an indirect way using a periphrastic such as 'make'.[9] This distinction is more generally used to predict argument selection, categorisation properties and word derivation.

They do note, importantly, that the external/internal distinction does not directly reflect real world facts, but our conception of them. Certain events can be seen as either or both type of cause. This is certainly the case and, for this author changes the focus of the work. (Levin & Rappaport-Hovav, 1995) does note that cross-linguistically certain verbs vary as to how the language construes them. But this also occurs within a language both diachronically and synchronically. This makes a prediction of the argument structure that is stronger than a default situation rather difficult.

For example, they note that *shudder* and *shake* are the same in every respect aside from the fact that the former is internally caused and thus cannot be used as a transitive causative. Shuddering is a certain type of shaking coupled with an emotion reaction. This emotional reaction is characteristic of a self-controlled body and usually it is only people, animals, and animate machines can be said to shudder. Things such as teacups and leaves can only shake. Thus *shudder* is an internally caused verb, with agency having nothing to do with the distinction.

However, while grammatical agency indeed has nothing to do with things, control certainly does. Take away the emotional component and *shudder* works fine as a transitive. Inanimate objects with no control or emotion can be said to shudder, presumably as the motion resembles an animate agent shuddering. This is from fan fiction found on Google.[10]

(6)   The warm wind shuddered the leaves of the old oak tree beyond the window.

The OED does have rare uses of 'shudder' as a causative verb as well. One use is

---

[9]They do note however, that some internally caused verbs can be used transitively but these sort of causatives represent a distinct phenomena.

[10]It may be the case that *shudder* and *shake* are not the same in every respect. There are many ways a tree can shake, but if it shakes in a certain way that (metaphorically speaking) resembles a certain type of shaking behaviour that a disconcerted person does when shuddering, it seems no stretch to extend its use to inanimates. Typically, the emotional component of shuddering is inferred from the shaking behaviour of the subject, and is not seen directly.

poetic, but does use 'shudder' in a sense that does include an emotional reaction in the causee. (From E. Blunden, *English Poems*, 1925)

(7)   A drowned sheep lodged In a black holt of alders, Its poor fleece brown and vile, To shudder young beholders.

Ultimately, it really is a matter of construal, along with context, information structure and a number of other factors that determine whether a verb can be used in a transitive causal construction. It is often the case that the linguist attempting to rule out certain uses is too constrictive.[11] For the interaction of various factors (including psychological construal) that influence the argument selection of various verbs see (Goldberg, 2000) and (McKoon & MacFarland, 2001).

The point of this is not to rubbish the *semantic* distinction between external and internal cause, but rather to change the perspective. Instead of trying to predict which sentences are felicitous or how a particular verb's argument selection is always like, the purpose is instead to predict the meaning if such an argument selection is used. So, rather than deal with such matters as predicting traditional thematic roles, argument linking and structure, it is better to examine the various ways something can be causal and to precisely formalise these. Should a form be used that is not conventional, I would rather explain how it works than pedantically star something that does not fit the theory at hand.

Nevertheless, the verb class distinction and distinction between external and internal cause is a good default starting point. But even when we allow that there can be a range of construal of a verb as internally or externally caused, they are most likely other factors that determine a verb's behavioural possibilities. For example, they correctly class certain verbs of manner of motion, such as *roll* or *swing* as externally caused when predicated of inanimate objects, e.g. *The ball rolled down the hill* or *The door swung open.* In their terms these are unaccusative verbs (where the subject is patient-like), that are restricted in their participation in certain constructions. For example, they are supposedly not allowed in reflexive resultative constructions or the 'way' construction.[12] Therefore

(8)   a. The door rolled open.
       b. *The door rolled itself open.
       c. The shutter swung shut.
       d. *The shutter swung itself shut.

(9)   a. *The pebbles rolled their way into the stream.
       b. *The ball bounced its way into the street.

---

[11](Levin & Rappaport-Hovav, 1995) does this as well with 'decay' and many other verbs that are supposedly only internal cause verbs but are often seen in their transitive counterpart.

[12]These constructions will be examined in more depth at the end of the chapter. For now, the data is enough.

Intuitively, the subject must be seen as *doing* something in the above. Given that the verbs are externally caused, it is rather the case that they are having something done to them, thus ruling them out of reflexive resultatives and the 'way' construction. Nevertheless, even when the context suggests that the motion is externally caused, it seems that their prediction is false, as can be seen from people's descriptions of bowling (taken from Google).

(10)   a. She bounced in anticipation as *her ball rolled its way down the alley*, leaping in victory as it struck down every pin in its wake.
        b. Frequent calls of "take two" could be heard on several courts as stray balls rolled their way into others' territory.[13]

Clearly in these examples the tennis and bowling balls' motion is externally caused (by the sportspeople), but nevertheless we can still view the balls (once set in motion) as doing something on their own, regardless of the external/internal distinction. The stereotypical motion associated with a shopping cart can also be used in the way construction[14] to give a poetic trajectory to its success.

(11)   The shopping cart thus rolled its way into marketing history.[15]

and even a glacier melting[16] can be conceive as a method of propulsion for the glacier, rather than something happening to it.

(12)   a. As the last great glacier slowly melted its way north, torrents of water and debris raced south.[17]
        b. When a retreating glacier – about 12,000 years ago – reached this soft seabed substance (as it melted its way north), the glacier was able to dig much, much deeper than when it dug into hard, non-porous rock.[18]

Similarly, inanimates can participate in the reflexive resultative construction as well.

(13)   a. Once the *telescope has swung itself into position and stops moving*, you'll see Saturn in the eyepiece.[19]
        b. The door *swung itself slowly open* with a languid creek, and he smiled as he went inside.[20]

---

[13] http://www.cavalierdaily.com/CVArticle.asp?ID=9973&pid=780
[14] This is a typical use of the 'way' construction, where the path generated by the verb needn't be literal motion, but a 'way' to a result, nonetheless.
[15] http://www.snopes.com/business/origins/shopcart.asp
[16] Sadly this will happen all the more.
[17] www.tomifobia.com/lowwater.html
[18] http://www.odysseymagazine.com/pages/askascientistarchives.php
[19] www.allenhayden.com/telescope.html
[20] users.tellurian.com/dano/story-argon.html

Certainly, the latter example is taken from fiction and one could argue that a bit of animacy is attributed to the door, giving it an internally caused construal. But the former is merely an instructional piece on how to use a telescope. Here (as with (13)), I believe that what is going on is that the external cause is ignored entirely. In these cases, the motion is started by an external cause, but (ignoring inertia, gravity and friction) once the motion starts it seems no great leap to construe the motion as now self-propelled and thus available for the constructions.

Finally, using a transitive structure with an internally caused verb such as 'walk' is normally considered to idiosyncratic and parallel to an extent with a periphrastic causal version. Think about the normal instances of 'John walked the dog' which is more than just ordinary causality but describes a rather complex event where the causee is really in control of the walking.[21] But when construed as the causee having no power, essentially the 'walking' is externally caused. As will be seen in the formalisation, this makes a huge difference and effects the event structure (imagine when someone is said to 'walk a puppet' rather than a dog).

What the above shows is that a deeper exploration into the notion of cause (in language) is necessary. One or two factors cannot account for the great variability of behaviour, matters are far too complex than that. Instead, we must look beyond the traditional formal semantic literature to arrive at a fuller notion of cause (or rather the multifarious types of causes). In the following sections we shall examine both in English and cross-linguistically what the various factors involved in causatives are, beginning with another semantic distinction.

Causal level

The notion of differing causal levels will be used to denote the scale of causation along the event pole, that is, how direct or indirect is a given causal event. Imagine the situation of ice melting where John purposely takes it out of the freezer and puts it in direct sunlight. It can be described both by

(14)   a. Sunshine melted the ice.
         b. John melted the ice.

(14-a) is a prime example of direct causation where the effect of the sun's rays and heat directly causes the ice to melt. It meets the criteria of no intervening events and temporal abutment. With the latter however, John could be on the other side of the world by the time the ice melts, and is a paradigmatic example of indirect causation. Both are expressed via the same form which may lead one to think that the type of causation is irrelevant. However, we could also express the situation as one of

---

[21](Levin & Rappaport-Hovav, 1995) note that these sort of transitives are a different phenomenon than the ordinary alternation. I agree, but when construed as a true causative, the logical implications do differ.

(15)   a. John let the ice melt.
         b. John had the ice melt.

But the above two sentences with 'sunshine' as the subject are a little bizarre.[22]
It may be objected that the sun's inanimacy is really the culprit in their infelicity.
However, using the famous example from Horn, we can see that causation, and
not animacy is really at issue here by using only animate participants. A classic
of example of the division of pragmatic labour involves two situations involving
Black Bart and the Sheriff. The first involves Black Bart shooting the sheriff and
the second involves sabotage of the sheriff's gun, causing it to backfire and killing
the sheriff. Horn proposes that in the following two sentences, only the first is
compatible with the direct situation and the second is only compatible with the
indirect situation:

(16)   a. Black Bart killed the Sheriff.
         b. Black Bart caused the Sheriff to die.

Conventional wisdom is completely correct in say that (16-b) is only compatible
(pragmatically) with the indirect situation, but (16-a) is actually compatible with
nearly any causal situation that has Black Bart as the first cause. Sticking to
the indirect causal situation, imagine the Sheriff's wife is out for revenge and ties
Black Bart to train tracks, awaiting the inevitable. Black Bart, not knowing who
she is, asks the wife why exactly she is doing such a dastardly act. Her natural
response would be

(17)   You killed my husband.

and less likely to be

(18)   You caused my husband to die.

yet this refers to an indirect situation. Once again the way the speaker views the
situation (not necessarily the situation itself) manifests in the choice of form used.
(17) removes the intermediate causal events and focusses strictly on the initial
cause. By using a direct causal structure she manages rhetorically to emphasise
Black Bart's culpability and not the gun's, in a sense making her revenge more
morally justified (on the Old West's morals at least).

   Here, it can be seen that for English, at least, granularity is one of the ways the
speaker construes the situation and the way it is coded linguistically is dependent
upon this and not necessarily relying on a speaker-independent notion of direct
or indirect causation. Granularity of this sort used in direct causal constructions
has aspectual implications as well that will be explored when we move on to
the formalisation. Cross-linguistically, this sort of construal may not be quite

---

[22]'Make', as Kemmer and Verhagen note, expresses and intermediate notion of causation that
makes it compatible with all sorts of causes.

so flexible, but there is a consistent coding relation between causal type and linguistic form used. We will first look at the range of causatives in English and then broaden our outlook.

5.2.2   English typology

Having examined the basic semantic concepts about causation and the way these are expressed in a number of languages, this section will focus on English causatives.[23] A connection to the basics of semantics of causation will also be made, but more specific causative quirks will be left until the closing sections, focussing on the formalisation.

English causatives

Observe the following well known pairs (or triples)

(19)   a. Sunshine melted the ice.
         b. The ice melted.
         c. John burned down the house.
         d. The house burned down.
         e. Fire burned the house down.
         f. Mary walked the dog.
         g. The dog walked.

The transitive sentences (*a* and *c*) imply the intransitive alternates. At first glance one could analyse (19-a) and (c) as ordinary accomplishments, with the intransitive versions as their passive alternates (e.g. *John built the house*) to *The house was built*. This however is not the case for a couple of reasons.

First, the (b) and (d) sentences are not passives at all, but, in fact what is called in the literature "anti-passives.". Compare the following:

(20)   a. The house was built (by John).
         b. The vase broke.
         c. The vase broke *by John.
         d. The vase was broken by John.

(20-a) is a typical passive, (b) a typical inchoative alternation and (d) a passive version of the active/causative *John broke the vase*. Typologically, something very interesting is happening here. While in both the passive (c) and intransitive sentence (b),[24] the verb form is different, as is the possibility for expressing an

---

[23]A brief examination of their diachronic evolution is presented near the end of the chapter.

[24]In the semantic literature, examples such as (20-b) would be called *unaccusative* and (19-g), unergative. Unfortunately, this terminology is almost exactly reversed in much of the syntactic and typological literature. To avoid this confusion, I shall use a variation of the terminology

agent. With the passive, the agent argument, though not obligatory, can be expressed using a 'by'-phrase. This is impossible with the intransitives. Sentence (20-c) is of the form 'S=P' but the event is presented as happening with no regard to an external cause. Croft notes that such constructions as the intransitive *break* are those in which "A is rendered nonsalient or nonexistent, and only P remains as the sole participant in the event." (Croft, 2001, p.317)

As touched upon, the verb form used by passives and the related intransitives is completely different. With a passive, a form of *be* is used along with a past-participle. With the intransitives, the verb is just the normal verb form that would be used with an active sentence. I used *break* in the last examples since it conveniently distinguishes morphologically the simple past and past participle, which is becoming rarer and rarer in actual usage.[25] The S=P intransitive sentences are a simple active intransitive similar to *John smiled* except the subject is now the patient instead of the agent.

S=P type intransitive forms are not the only type of intransitives that can be related to a transitive causative form. The pair in (19-f) and (g) can also be seen as a transitive 'alternation' but (20-g), is not of the form S=P, but rather S=A, that is the dog does the walking and is not undergoing something in the way ice melting does. But, regardless of the argument status, the transitive versions of both are lexical causatives. Note that this sort of agentive cause patterning is not productive, but can be a matter of historical accident or convention. For example, *walk* (in the literal sense) occurs quite easily in this alternation, but a verb of the same type, *run* can usually only be used causatively in a more metaphorical sense, e.g. *running a business*, or in a resultative, e.g. *run someone into the ground*. Someone who takes their dog for a run in the park, cannot conventionally be said to have 'run their dog'.

It is important to realise that the fact that a given verb does not participate in this transitive/intransitive alternation does not mean that there is no causation going on. For example, it would be rather hard to deny inherent causality in 'Mary cut the cake'. Yet, unlike, *break* or *melt*, one cannot say 'the cake cut'.[26] Instead one should view these causatives as a special type of grammatical construction, that presents causation in a certain way. Aside from certain syntactic restrictions, I have seen no account that can correlate the nature of the verb with its participation in the alternation absolutely, though there is indeed a degree of patterning by verb class.

---

of (Dixon, 2005) and (Croft, 2001). 'S=A' means that for an intransitive like *John walks*, the subject, John is agentive in nature. 'S=P' means that for an intransitive like *the ice melted* means that the subject is patient-like, i.e. being affected by the melting.

[25]In fact, one is just as likely to hear *The vase was broke* – for all I know this is not a recent phenomenon and indicator of degraded language ability. There is perhaps a continuing trend to conflate past tense and past participle.

[26]For this sections, middle voice is ignored, but something like 'the cake cut easily' is perfectly acceptable.

The syntactic restriction is that a transitive sentence cannot be used in a lexical causative situation. With any lexical causative, the subject corresponds to the causer of the verb, e.g. if Mary broke the vase, she was the causer of the breaking. But, now begin with *John cut the cake*. Supposing he did this under Mary's instruction, it must be expressed as an analytic causative, such as *Mary had John cut the cake* (contrast with the more morphological and case marking Turkish ((56))). Trying to express this non-periphrastically, we get 'Mary cut John the cake' or 'Mary cut the cake PREP John'. In both cases, Mary is the one doing the cutting, which is not the causality desired. The transitive, already having a direct object makes it rather difficult for the causee to also be a direct object. Thus, the alternation from S=A intransitives to a causative like 'Mary walked the dog' may indeed be acceptable. Note also, that my gloss of (45) ('I coughed the child'), while conventionally unacceptable is perfectly understandable nevertheless, as there is no syntactic restriction here. See (Dixon, 2005) for more details.[27]

Finally, aside from the lexical causatives examined in this section, there is a more productive way of expressing causation, the analytic causative. There are many auxiliary verb choices involved, and the next section we shall compare different analytic causatives to each other as well as to lexical causative to uncover the semantic distinctions already made.

Different English causative types and semantics

From (Dixon, 2005, p. 312)

(21)   a. John walked the dog in the park.
       b. John made the dog walk in the park.

(22)   a. Mary opened the door.
       b. Mary made the door open.

(23)   a. Fred dissolved the sugar in the liquid.
       b. Fred made the sugar dissolve in the liquid.

The (a) sentences all imply that the causal event went in the expected manner, and will also be associated with direct causation (either manipulative or directive), while the (b) sentences can be used for either indirect causation or to imply that the causal event was not carried out in the normal manner. For example, (21)(a) causes one to imagine that John merely took the dog in the park and walked with him. Sentence (21)(b) can also be direct causation, but conjures up some use of force by John to get the dog to walk. This could still be direct causation, but

---

[27]One could also try to come up with such alternates as 'Mary John cut the cake', but these are mere nonsense.

the force relation between the participants is highlighted.[28] (23)(a) and (23)(b) are periphrastic causatives of an S=A type, and one imagines that with (23)(a), Fred needed to do more than merely put the sugar in the liquid, but had to resort to additional means. Depending on the means, this could then be an example of direct or indirect causation, but, either way something other than the normal way of dissolving sugar is going on.

Again, semantic differences can arise along a lexical/analytic scale. As seen with many other languages, the difference in the construal of closeness of cause can lead the analytic forms to have a variety of other implications (including difficulty of realising the caused event). Remember, however, that even if Fred needs to dissolve the sugar in a roundabout, complicated way, (23)(a) is still applicable. Should the lexical causative be a construction that means not only causality, but that the causality is done in a direct or normal way, this does seem a bit odd. But, if one is only concerned with Fred being the cause and the sugar dissolving, the lexical causative is perfectly acceptable – one merely decides not to present the indirectness as it is irrelevant to what the speaker desires to say. As we shall see later, the notion of directness seems to have expanded allowed implications to arise to the subjects intention or responsibility, and this is often enough to make recourse to a lexical causative perfectly fine in these circumstances. Should the speaker want to emphasise the difficulty or novel way Fred achieved his goal, then (23)(b) is the appropriate form.

The above contrasts the use of lexical vs. periphrastic causative types, but it has already been established that if one wants to make a causative of transitive sentences, then the only way to do this is by using an analytic causative. Therefore, any putative distinctions between the level of causation or relation between the participants cannot be elucidated through contrasting syntactic forms. Also, unlike the languages studied in the cross-linguistic section, there is no recourse to differential case marking to make these distinctions. Instead, the choice of auxiliary verb becomes all important.

(24)   a. Mary made John cut the cake.
        b. Mary had John cut the cake.
        c. Mary let John cut the cake.

(24-a) would be used to express directness of cause, as well as a degree of force involved on John. For example, it could be both direct and manipulative, where Mary actually manipulates John's limbs to get the cake cut. It could also be indirect, but directive, involving Mary commanding John to cut the cake. (24-b) on the other hand implies a degree of indirectness, and can express a situation where Mary is organising a party and assigns various tasks via a list or an assistant. It is also much easier to imagine with (24-b), that John was quite willing to

---

[28]One could also with this sentence imagine various bizarre ways of indirect causation used to get the dog to walk.

cut the cake (or at least more willing than in (24-a)). (24-c) is a typical example of a permissive construction. The reader may also want to revisit (4) for more semantic differences between various analytic types.

Verbs such as *force*, *get to*, *cause* can also be used to express the causal relations between the events or the participants. The perceptive reader will note that with any of these analytic constructions, the causal distinctions are not nearly so clear-cut as will be seen in the cross-linguistic section. Certainly, *force* is clear, but depending on the verb and situation, *make* can be either direct or indirect and can even be stretched to the situation where Mary uses her assistant to tell John to cut the cake. But, the use of *make* does imply that even though Mary accomplished the causation indirectly, she was wilfully using her power to force the causal event (regardless of the causee's wishes). This isn't the case with *have* or certainly *let*. But, if we look at these causatives as not just directly constraining the 'real' situation, but as also constraining the way the situation is presented, this apparent difficulty dissolves. We shall now see that this sort of phenomenon is exactly what happens with lexical causatives as well.

Lexical causatives and level of cause

It is typically assumed that lexical causatives are associated with a direct cause and the corresponding analytics with an indirect cause. However, actual level of causation has nothing to do with whether the causee can be the subject of a lexical causative

(25)   a. Fire burned down the house.
        b. John burned down the house.
        c. The landlord had John burn down the house.
        d. The landlord burned down the house.

(25-a) is an unequivocal example of direct causation. (25-b) in the situation where John pours petrol in the house and lights a match is also normally considered to be an example of direct causation, even though one could make recourse to the petrol and match sub-events. (25-c) is an analytic, where the landlord, wanting to evict the current tenants and rebuild expensive yuppie flats is the ultimate cause of the fire. But he is not seen as a direct cause as he was most likely temporally unconnected with the burning event and the intermediary events are clear. Nevertheless, (25-d) is also a perfectly acceptable way of construing the situation, and both types are regularly used depending on the situation and the speakers intentions. Indeed, in regard to the landlord, it can be seen as a much stronger statement.

The example of Black Bart and the sheriff already touched upon this point. Whether the 'causer' availed himself of a direct or indirect causal event, the lexical causative rhetorically makes the subject more responsible by using a form associated with directness of cause. Lexical items often change diachronically to

expand or change their meaning in regular ways,[29] and this seems to be exactly what happens with at least some English causatives. But it is not the case of it only being a lexical item expanding its meaning or domain of application, but rather also the construction itself.[30]

5.2.3   Cross-linguistic typology

In the previous section we examined the semantic types of causality and saw that the matter is much more complicated than a simple notion of cause and effect. Both the relationship between the events and between the participants give rise to a variety of causal situations. The morpho-syntactic coding of these relations is the topic of this section, as is their relationship to the different semantic types. Be advised that this section's intention is not to give a complete typological account of the world's languages in this area, but only to allow the reader to see both the variation and similarities among the worlds languages, and how their coding directly parallels the informal semantics sketched above. Many issues such as Comrie's case and animacy hierarchy are ignored (Comrie, 1989). For a more elaborate explanation of a typology of causatives, the reader is encouraged to look at (amongst others) (Comrie, 1989), (Comrie, 1985a), and (Song, 2001).

This section may initially cause bafflement as to its relevance to the subject. But, by beginning with examining the various ways verbs in various languages change valency is the first step in understanding how causatives work. Taking seriously (Kemmer & Verhagen, 1994)'s proposal that non-lexical causatives are related to 'ordinary' transitives, it is helpful to first look at ordinary verbs, their morphological and case-marking effects and then look at parallels with causatives. The same morpho-syntactic mechanisms used in valency change are used in causatives, with markedly different semantic effects however. This helps determine the constraints and properties of linguistic causation, a necessary step to any formalisation.

Valency Change

As is well known in English, both ordinary verbs and causatives allow a change in valency (number and role of syntactic arguments) with no change to the verb or case-marking whatsoever. For example (Comrie, 1985a):[31]

(26)   a. Mary is reading a book.
          b. Mary is reading.

---

[29]See e.g. (Traugott & Dasher, 2002).

[30](Kemmer & Verhagen, 1994) posit a similar effect with the languages where case marking normally indicating indirectness and independence on the part of the causee, can be used to mark topicality of the caused object.

[31]Unless otherwise noted, all examples in this section come from (Comrie, 1985a).

      c. John is boiling the water.
      d. The water is boiling.

The form of the verb in both sets of examples does not change, yet especially in
the *boiling* example, the valency changes radically. In (26-c), the affected partici-
pant is the water and is syntactically the direct object, while in (26-d), the water,
while retaining the same semantic role is now syntactically the subject.[32]

    In many languages, ordinary transitives require special verbal marking in order
to express the rather simple valency difference between (26-a) and (b). Wolof is
an example where the reduction of valency requires no verbal change, but the
addition of an indirect object where the verb is not by default ditransitive does
require one. In other words, adding an argument on a verb that normally does
not allow that many arguments requires the verb to change form.

(27)   Mungi   dyàng tééré bi
          PRES(3SG) read   book the
          He is reading the book.

(28)   Mungi   dyàng
          PRES(3SG) read
          He is reading.

(29)   Di   na   nyu la dyaay suma xar
          FUT AUX 3PL you sell   my   sheep
          They will sell my sheep to you.

(30)   Di   na   nyu dyaay suma xar
          FUT AUX 3PL sell   my   sheep
          They will sell my sheep.

(31)   Mungi   dyàng-al eleew yi   tééré-èm
          PRES(3SG) read-   pupil the(PL) book-his
          He is reading the book to his pupils.

The first pair (27) and (28) directly parallel (26)(a) and (26)(b), where the direct
object is merely dropped. Similarly, the morpho-syntactic difference between (29)
and (30) involves merely dropping the indirect object, where the verb (*sell*) allows

---

[32]Of course there are suppletive causal pairs such as *fall/fell* and even *die/kill* where different
verbs are used depending on whether a transitive of intransitive form is used. Being suppletive,
this relationship is not productive, but rather unpredictable and exceptional.

ditransitivity as a basic valency. But when a basically transitive verb requires the addition of an indirect object, the verb requires a suffix in order to allow this.

This sort of patterning is rather common – languages such as Swahili (and other Bantu languages) use an applicative suffix on the verb to add an extra argument to the verb (with some object order change), while Russian requires a the verb to change in the case of valency reduction for some verbs[33]

(32)   Boris čitajet knigi
        Boris reads   books.

(33)   Boris čitajet.
        Boris reads.

(34)   Kury nesut jajca
        Hens lay   eggs.

(35)   Kury nesut-sja
        Hens lay.

The first pair, (32) and (33) show that the verb *čitat'* requires no derivational morphology change to omit its direct object while the second pair, (34) and (35), shows that the verb *nesti* requires the verb to take the reflexive suffix *-sja* (or *-s'* after vowels) to reduce its valency.[34] Hungarian, Swedish, Dyirbal and Georgian all have a similar type of derivational morphology, with it being the case in Dyirbal that it is systematic enough that this sort of derived passive is often named the 'antipassive' voice.

Only a few of the possibilities of valency change have been sketched and the complication of the derivational morphology affecting both case marking and aspect has not been shown. Nevertheless, one may wonder what any of this has to do with causatives.[35] Recall the so called causative alternation is not merely a valency change, but really a change of a core meaning of the verb – its argument structure. To repeat an oft used example, the difference between

(36)   a. Sunshine melted the ice.
        b. The ice melted.

is not just of valency, but the fact the subject in (36-a) is the causer and the subject in (36-b) is the causee. As with ordinary English verbs, there is no

---

[33]The examples below refer to habitual situations. For some of the verbs that change in habitual aspect, they do not change form in the case of a sentence referring to a specific instance.

[34]Could there be a patterning according to aktionsart? Unfortunately two examples is not enough to say anything interesting.

[35](Comrie, 1976b) examines valency change and related case marking as a purely syntactic phenomenon. That is, as verbs take additional arguments the cases of the various objects and direct objects alternate in a hierarchical pattern. (Song, 2001), (Kemmer & Verhagen, 1994) and even later Comrie view case marking in many languages as expressing semantic differences in causation as well.

change in verb form when this happens, but a number of languages do need to mark the verb when there is a valency change that changes the thematic nature of the subject.

The Causative Scale – bringing periphrastics into the picture

[36]

Examine the following three sentences from Nivkh (Slavic language, productive morphological causative is *-gu*)

(37) Lep   če-d'
      bread dry(INTRANS)-
      The bread dried.

(38) If   lep   seu-d'
      He bread dry(TRANS)-
      He dried the bread.

(39) If   lep   če-gu-d'
      he bread dry(INTRANS)-cause-
      He caused the bread to dry (e.g. by forgetting to cover it)

(38) has a simple transitive causative that (when comparing to the verb in (37)) can be considered lexicalised in the same way that suppletive pairs (e.g. *fall/fell*) are in English. But in (39), the intransitive form has a causal marker converting the verb from intransitive to transitive. Both (38) and (39) are transitive, yet as can be seen from the glosses, their meaning does differ as far as the type of cause. (38) implies that the subject did something directly (and perhaps intentionally) to dry the bread, while Comrie's gloss of (39) makes clear that the causation at stake is indirect causation, here a possible situation is merely leaving the bread to dry by accident.

In Nivkh, the contrast is between a lexical and morphological causative, where the situation expressed by the lexical causative implies direct causation, whereas the latter expresses indirect causation. Russian on the other hand, has lexical causatives (40) as well as morphological causative markers (*-sla* or *-s*), and yet another way to express causation

(40) Anton slomal paločku
      Anton broke the   stick.

(41) Anton sdelal   tak, čtoby paločka slomala-s'
      Anton brought it   about that   the   stick broke.

---

[36]These type are also called analytic or syntactic, and the terms will be used interchangeably.

where the latter (as is easily seen in the gloss) directly parallels the English, in that
the lexical version implies direct causation, and the latter indirect. (41) makes
clear that the strict morpho-syntactic distinction between lexical, morphological
and analytic causatives is not always so clear cut. Unlike the simple, lexical
transitive of (40), the latter example combines both periphrasis and morphological
causation (*slomala-s* is an anti-causative with a morphological suffix to derive it).
Therefore, the causativity of the causing event in (41) can be unequivocally seen
as an analytic causative. However, what appears to be strict analytic causation
in French is not quite what it seems at first.

(42)   Paul fit   manger les pommes à/par Pierre
         Paul made to-eat   the apples   to/by Pierre
         Paul made Pierre eat the apples.

Faire is used along with the basic verb, working in French similarly to *make*
or *cause*, and appears to be the equivalent of the English analytic construction.
Thus, *faire courir* is 'to make run' and *faire manger* is 'to make eat'. But, a
closer examination shows that the English parallel does not hold

(43)   *Paul fit   Pierre manger les Pommes.
         Paul   made Pierre eat   the apples.

(44)   Paul lui   /?*le fit   manger les Pommes.
         Paul to-him /him made to-eat   the apples.
         Paul made him eat the apples.

    Comrie notes that a single French predicate cannot take two subjects or two
direct objects. In (43), both Paul and Pierre are subjects, and in the English
gloss, Paul is the subject of *made* and Pierre the subject of *eat*. Therefore one
would expect that (43) should be grammatical as it conforms to this rule as while
there are two subjects, there appears to be only one subject per verb. However,
faire with an infinitive behaves in many cases as a single verb, making the like of
(43) impossible. In (44), 'him' must be in the dative case as 'apples' is already
a direct object. Therefore, *faire + inf.* can really be considered as either an
analytic or morphological morphological causative, depending on one's analysis.
We have already seen a similar difficulty in sometimes distinguishing whether a
particular causative is lexical or morphological such as the existence of suppletive
pairs, and not fully productive morphological causatives.[37]
    Nevertheless, we can still imagine a morpho-syntactic scale of causes with two

---

[37]For example, in Russian, we saw how the causative suffix is only applicable to some verbs
with no discernible syntactic or semantic patterning. In Turkish, it is possible to form a
causative with nearly any verb. Causatives in Turkish are unique enough to get a special
examination at the end of the cross-linguistic section.

endpoints – *lexical* and *analytic* with *morphological* in between. Corresponding to this syntactic scale, is the 'scale' with endpoints of direct and indirect causation. Lexical causatives with correspond with direct causation situations (or the interpretation of the situation), while indirect causal situations correlate with the analytic type – in languages where more than one type of causative is available. This shall be further elaborated upon in 5.2.4, but first recall that the difference between direct and indirect causation was not the only semantic distinction made in (4).

Case Marking

The last section dealt primarily with different expressions for either direct or indirect causation and the continuum from lexical to analytic causatives. This involves different ways of linguistic coding to express the relations between the causing and causal event. Differences in case marking, can express the other main causal distinction, that between the causer and the causee. The following is from Hungarian. (Comrie, 1989)

(45)   Én köhög-tet-te-m   a   gyerek-et
         I   cough-CAUSE-PAST-1SG the child-DO
         I coughed the child.[38]

(46)   Én köhög-tet-te-m   a   gyerek-kel(INSTR)
         I   cough-CAUSE-PAST-1SG the child-INST
         I made the child cough.

Both of the above can have a direct causal interpretation, where there is no intervening event between the subjects action and the child coughing. In fact, both sentences are examples of morphological causatives, so the same causal interpretation on the event level is no surprise. The formal difference between the two examples is that in (45) the causee (the child) is marked as a direct object (accusative case) while in (46), is with an instrumental. Their is also a semantic difference. In (45), the causee retains little sense of control and suggests a situation where the subject slapped the child on its back. (46) is more appropriate where the child is merely told or asked (perhaps by a doctor) to cough. Both are examples of direct causation, but do express the distinction between manipulative and directive causation.

Generally, there is said to be a patterning that the more the causee has control of the caused situation, the more likely it is to be marked in a non-accusative case. (Cole, 1983) suggests a scale:

---

[38]This gloss is not allowed in English, but its use in the direct, lexical construction allows the gloss to express the manipulative causal meaning in the Hungarian.

(47)   Instrumental > Dative > Accusative

where the nature of the causee is more agentive on the left end, exerting more
control, while on the right end the causee is more patient-like, and exerts less
control of the situation. It is then easy to how markers associated with greater
and lesser control by the causee can be used to express the difference between
manipulative and directive causation. Direct physical manipulation of the causee
would usually mean that the causee has little control over what happens, while
non-physical causation may allow the causee some degree of freedom as to whether
perform the caused event. However, Song's weaker distinction is more correct –
the causee may be in a non-physical causal situation that but still be affected
enough by the causer to have little choice in carrying out the action. From
Bolivian Quechua (Cole, 1983):

(48)   nuqa Fan-ta   rumi-ta   apa-či-ni
          I   Juan-ACC rock-ACC carry-CS-1SG
          I made Juan carry the rock.

(49)   nuqa Fan-wan   rumi-ta   apa-či-ni
          I   Juan-INST rock-ACC carry-CS-1SG
          I had Juan carry the rock.

In neither of the two examples is there manipulative causation, but in (48), Juan
is perceived to have no control as to whether or not he will carry the rock (perhaps
he is in a prison chain gang). Sentence (49) works better in a situation where Juan
submits voluntarily to an order or request. The glosses indicate that in English,
the degree of affectedness of the causee may be made by the use of different lexical
auxilliaries. Case marking also affects topicality and the goal of the causer. From
Hindi (Kemmer & Verhagen, 1994):

(50)   Mai-nee raam-koo kitaab park-vaa-ii
          I-AGT   Ram-DAT book   read-CAUS-PAST(f.)
          I had Ram read the book.

(51)   Mai-nee raam-see   kitaab park-vaa-ii
          I-AGT   Ram-INST book   read-CAUS-PAST(f.)
          I had the book read by Ram.

Sentence (50) indicates that the causee's goal was to get Ram to read the book,
but in (51) Ram is not topical, but the aim is simply to have the book read.
According to the scale in (47), with (50), the causee is more directly affected
(in comparison to instrumental marking) than the causee in (51). According to

(Kemmer & Verhagen, 1994), "the more affected the participant, the more object-like and therefore central...to the event as a whole" (p.133), while the causee with instrumental marking is a more peripheral participant. The authors note that "periphality normally would not cohere with topicality" (p.133). Sentence (51), the more peripheral causee of the pair, fits less well as a topic (in comparison to (50)), thus allowing the book to be the topic of the sentence.

Their generalisation is that where a language admits the two oblique cases (dative and instrumental) to be used in contrast, there is the same semantic patterning. (Kemmer & Verhagen, 1994, p. 135)

> DATIVE = more integrated, highly topical participant, affected as a human entity vs. INSTRUMENTAL = low degree of integration into event, low topicality, low degree of affectedness.

In a similar manner to the distinctions between types of causatives, there is a scale regarding the case marking of causees (depending on the contrasted case markers) that reflects both the relationship between the cause and the causee, as well as how affected and integrated the causee is to the event as a whole. Agency and periphery of the causee lie on one end of the scale – the polar end having the causee be more patient-like, with less control over the caused event.

5.2.4   Causative types and semantics

Thus far, a case has been made for distinctions in the notion of causality in regard to direct or indirect causation between the causal sub-events and level of control and independence between the causer and the causee. The last two sections have illustrated how these various notions of causality are coded by different languages. (Song, 2001)

(52)   lexical   morphological   syntactic [39]

For Song, "there is a strong tendency for manipulative or direct causation to be mapped onto the causative types at the left of the continuum in preference to those on the right...Directive or indirect causation...is far more likely to be expressed by the causative types to the right of the continuum." (p. 278). This introduces a slight complexity to what has be seen so far. (52) has previously been used only in pairs to point out the differences between direct and indirect causation, but there is also a correlation with the distinction between manipulative and direct causation. Mixtec (Song, 2001)

(53)   sáʔà hà nà   kee
         cause NR OPT eat
         Make him eat.

---

[39]Song's use of 'syntactic' is equivalent to the use of 'analytic' in preceding sections.

(54)   s-kée
          CP-eat
          Feed him.

The causal differences can be partially brought out by examining the English glosses. Both are examples of direct causation, but (53) implies that the causer induced the causee to eat by suggesting (or commanding) that the causee eat the food available. Sentence (54) is more appropriate where the causee is either refuses to eat food or is unable to (perhaps in a coma). The morpho-syntactic difference is a contrast between an analytic causative and a morphological causative, while the semantic contrast is between manipulative and directive causation, rather a direct/indirect distinction.

In fact, it might be better to change the English glosses to something more parallel to the situations described: (53) could be glossed as *Feed him*, while the situation given for (54) would be perfectly appropriate with the lexical *Force-feed him*. Both, then, are lexical causatives, consistent with the notion that both situations are examples of direct causation.[40] Song gives a number of similar examples from Japanese and Newari where the formal difference between morphological and analytical causation is also used form the directive/manipulative distinction. (p. 279-280).

Both scales, (52) and

(55)   Instrumental > Dative > Accusative (=(47))

correlate morpho-syntactic marking with a degree of closeness, either closeness of the causal events or of the participants associated with the events. For the first scale (52), Song appeals to a notion of iconicity or iconic distance to explain the correlations along the from the left to the right of the scale. Quoting (Croft, 1995), iconic distance can be seen as "the principle that the structure of language should, as closely as possible, reflect the structure of experience expressed by language" Space does not permit here but the reader is encouraged to see (Croft, 1995) and (Song, 2001) for a description and analysis of the relations between semantic distance and conceptual distance. (Song, 2001) takes this a bit further, attempting to construct a cognitive notion of causation, along with a semantic typology of causation.

---

[40]Recall that for English, *make* is a sort of causal middle-ground and can be associated with either direct or indirect causation, and with *eat* can be seen as sometimes synonymous with 'feed'. Simply leaving food and the permission (or lack of prohibition) to eat it is less easy to call *feeding*. The UK has an expression to keep someone *fed and watered*. This quasi-passive construction can be related to the use of differential case marking to indicate the degree of control of the causee. Notice that the formal difference remaining (of the new glosses) only involves the choice of verb used, with *force-feeding* not only meaning direct causation, but also implying a relatively low degree of control on the part of the causee.

The second scale (47) is used by (Kemmer & Verhagen, 1994) to further their proposal that transitive causatives[41] are modelled upon simple transitive structures. Should the causee be directly affected by the causer's action, where differential case marking is available, the causee will be marked by the accusative case, just as a patient is marked in an ordinary transitive cause. Recall that dative marking may be used (in opposition to accusative) to indicate that the causee is somewhat in control of the causal situation. Transitive, experiential verbs in Bolivian Quechua (there is a also a causative suffix) also exposes this difference:

(56)   nuqa runa-man rikhu-či-ni
        I   man-DAT see-CS-1SG
        I showed it to the man.

While the experiencer is shown something, note that the verb is really something like 'cause-to-see', but unlike being hit with a hammer, the experiencer does have some choice as whether to see it or not.[42] Instrumental marking separates the causee from the causer even further. For a further elaboration on this notion as well as a Dutch corpus study showing that differences in the causee's agency depends on both the verb used (*laten* or *doen*) along with how the causee is marked, the reader is encouraged to see (Kemmer & Verhagen, 1994).

To conclude this excursion into the world's languages, it is important to remind the reader of the reason for the detour. Our examination of causality began with purely semantic distinctions encouraged by examining a few sentences of English along with various interpretations or imagined situations. Often the same lexical item or syntactic form could be used somewhat ambiguously, that is it could apply to various types of causal situations. Languages that make these distinctions more explicit are useful for strengthening the notion, say, that gradations of the causee's control is not merely an effect of different ways of envisioning the situation, but is encoding it (where it is morpho-syntactically possible) formally into the utterance. Thus, any semantics of causation will need to take these notions seriously and come up with ways of accounting for all of them.

To be sure, some of the differentiation is purely syntactic such as the French examples in (42), (43) and (44). In English, syntactic reasons alone account for the fact that transitive lexical causatives can only apply when the causee is patient-like or the agent of an intransitive – that is, many transitive verbs in English lack a simple causative use (see 5.2.2) due to purely syntactic reasons. Nevertheless, we have also shown how both semantic and cognitive notions can used to account for various morpho-syntactic phenomena, and can be useful in

---

[41]They explicitly ignore lexical causatives in their paper.

[42]Chapter 4 on perception verbs incorporates the perceiver's attention or readiness to see directly in the formalism.

accounting both for developing our semantic account as well as examining the development of language diachronically. But before moving on to English, a final excursion into Turkish causatives makes clear the notion of an extended causal chain, as well as providing an example where the case marking differences are strictly syntactic, with changes to the verb doing the semantic work.

Turkish holiday

Our last look at causative typology is a quick look at the way Turkish allows causatives of ditranstives, where a rather complex situation can be expressed with only one verb, and clever use of case marking.[43] We have already seen the way various languages can use a causative prefix or suffix along with a change in case marking to turn a mono-transitive sentence into a causal one. In Turkish, it looks like

(57)   Müdür mektub-u imzala-di
         director letter-ACC sign-PST
         The director signed the letter.

(58)   Disci   mektub-u Müdür-e   imzala-t-ti.
         dentist letter-ACC director-DAT sign-CS-PST
         The dentist got the director to sign the letter.

Comrie proposes a hierarchy of grammatical relations to reflect the status of the participants.

subject > direct object > indirect object > oblique object

The encoding of the causee is that of the leftmost unfilled position on the hierarchy. Thus, in (58), the causee (the director) must be marked as an indirect object, as the subject and direct object slots are taking up by the dentist and the letter, respectively. While even such a construction such as (58) would be impossible in a morphologically impoverished language such as English, things become even more remarkable when we look at what happens with ditransitive sentences.

(59)   Müdür Hasan-a   mektub-u göster-di
         director Hasan-DAT letter-ACC show-PST
         The director showed the letter to Hasan.

---

[43]All examples in this section are from (Comrie, 1976b) and (Comrie, 1985a).

(60) Disci Hasan-a mektub-u Müdür tarafindan göster-t-ti
dentist Hasan-DAT letter-ACC director by show-CS-PST
The dentist got the director to show the letter to Hassan.

In the standard ditranstive the director is an ordinary subject, but when turned into a causee in (59), he must now be encoded as an oblique, as the higher-up slots are already filled by the causer, the letter and Hassan. Unfortunately, as is pointed out in (Song, 2001), Comrie's hierarchy appears not to be universal and exhibits some exceptions in the few languages that mark causatives in this particular fashion.[44] Nevertheless, the examples from Turkish suffice for our purposes.

First, it is (for readers mostly familiar with only Germanic or Romance languages) a quite mind-boggling way of expressing quite complex causal situations. Second, it shows the importance to the casual scenario of the status and relationship of the participants.

5.2.5 Participants and Causal Level – steps toward English formalisation

Having examined various types of linguistic causality, it is now time to connect these more concretely to their temporal and causal implications. Here we will see how the role of the causer and causee participants and their relationship to each other can have a profound effect on the implications of the discourse, especially when looked at in varying tense and aspectual configurations.

Notion of cause revisited – the implications

In our earlier exploration of English, it was noted that the following pair of sentences do differ. While they are both used in a direct, lexical causative construction, they have slightly different properties.

(61) a. Fire is burning the house down.
      b. John is burning the house down.

Sentence (61-a) closely resembles an ordinary transitive and, indeed, behaves the same as an ordinary activity or accomplishment. For example, the use of the non-futurate progressive implies that the house was indeed on fire. Removing the telic 'down', the imperfective paradox vanishes and one could conclude that 'fire burned the house'.

The causal chain takes a step backward for (61-b), which while direct in form, focusses on an animate agent that is somehow utilises an elemental cause to accomplish the goal. What is important is that (61-b) has no implications

---

[44]Comrie himself notices this, as well.

whatsoever. This may sound odd and an actual example may help. It is taken from the TV series Arrested Development and involves a conflict between father and son. The son is working at the family banana stand, but for reasons of making a rebellious statement decides to burn it down. He is busily putting newspapers at the base of it when his father arrives. Note that there is no fire at this point. The father is Michael and the son, George Michael.

> *Michael*: (approaches) George Michael?
> *George Michael*: (stands quickly) Dad!
> *Michael*: Hey, what are you doing?
> *George Michael*: (*tries to think of something*) I was just, uh... (...but quickly gives up) I was burning down the banana stand.

Should 'fire' have replaced 'I' in the last sentence, it would be false as the banana stand had yet to burn.[45]

When accounting for elemental verbs such as 'burn' or 'melt' the level of cause associated with the casual participant must be taken into account. The most direct level of cause is when the causee is naturally associated with the process denoted by the verb.[46] Here there is quite a direct link and one assumes a temporal parallel between the fire and the act of burning – something that isn't the case with a personal agent.

The agentive version really has a double telicity in the progressive – the telicity of getting the house burning and the telicity of the house burning down. We will see this again in the next section that focusses on relations between participants. Formally, the elemental and personal agency cause will need to be modelled slightly differently from each to account for the differences in regard to the implications.

Desiderata of participants and events/eventtypes

The last section showed the implicational differences between a direct lexical cause form used for both truly direct (elemental) and quasi-direct (personal agency) types of causal events. This section will show that the relation between the participants which some languages mark morpho-syntactically also have implicational differences. Again, by examining the same causal events and participants in different aspectual and tense configurations. light can be shed on the differences and how to handle them.

(62)   a. The gardener is clearing the leaves.
         b. I've had/made the gardener clear the leaves.

---

[45]It does do so later in a deliciously ironic way.

[46]*Fire froze the football pitch* is not such an example, as while the causee is an element, it is the wrong type. In this case fire plays a sort of personal agent role – say fire destroyed the under-soil heating on a cold night.

As the caused event of (62-b) is a transitive (where S=A), there is no lexical way to express my causal effect upon the gardener. Instead, a periphrastic must be used, and in normal circumstances there is no implicational difference between *have* and *make* as far as the events are considered.[47] Assuming only a directive causation scenario, the gardener obeys the orders and is the one ultimately in control of the leaf clearing.

Imagine that I say (62-b) while the gardener is indeed in the act (62-a). I can say to an inquisitive neighbour either (62-b) or the progressive version (*I'm making him clear the leaves*). This much is obvious, but the interesting part is that the gardener need not finish clearing the leaves and my utterance of (62-b) is still perfectly felicitous. From my causal point of view, my job is done if I manage to get him started on the job. The gardener could get fed up or be hit by lightning, yet my causal duty has been completed. Telicity as far as the causer is concerned only applies to the causing of the gardener.

Contrast this with the earlier scenario of John and the doomed house. The parallel *John has burned down the house* or even the periphrastic version *John has made the house burn down* has a rather different implication. Both of these utterances[48] are only compatible with situations that the house burned down. Here telicity goes beyond the causal event starting a caused event, but extends completely to require the caused event (if telic) to run to a natural completion.

At first glance the difference between the two caused situations seems quite clear cut. *Burn* is an S=P type or what would also be called an externally caused eventuality, while the sense of *clear* used here is an S=A type, or an internally caused eventuality. However, the related notion of control (which usually correlates with the internal/external difference) is actually the culprit here.

(63)   a. I've walked Mary to school.
         b. I('ve) had/made Mary walk to school.
         c. I made the puppet walk.
         d. I('ve) made the puppet walk across the stage.

All of the above examples use the internally caused (S=A) *walk*, but the type of participants and context change the result of applying the 'perfect' test used above. Take (63-a) and (b) as expressing typical examples of directive causation. The direct causal version where I walk along with Mary (perhaps she is a small child who needs help) implies that Mary walked to school. Should I leave on her own halfway there, (63-a) no longer seems to be a proper utterance, but (63-b) is. (63-b) is rather similar to (62)(b), in that convincing Mary to walk to school is enough. Once she leaves for school, I've done my causal job and (b) is perfectly

---

[47]There is a slight coercive difference between the two. The former does suggest that I merely asked the gardener while the latter suggests something a bit stronger. However, sticking to a directive causation scenario this makes no difference for our purposes.

[48]This also holds for the simple past.

felicitous.

(64)   I never should have made Mary walk to school, but I wasn't aware of the
        kidnapping gang in the city.

With this I successfully made Mary walk to school (either relent or in this case
actually begin walking) but there is no guarantee that she makes it to school.
As far as the complement is concerned the bare form version of the imperfective
paradox still holds.

Sentences (63-c) and (d) use the same caused verb (*walk*), but a puppet has
far less control than a human being, and can only be walked by direct control,
i.e. manipulative causation.[49] Thus (63-d) is only felicitous if the puppet makes
it from one end of the stage to the other with my help. The rather odd scenario
where I drag Mary down the street and treat her as if she were a puppet has the
same implications. Importantly, in this scenario Mary has no control over here
walking, and the implications from the perfect are the same as for the puppet
walking or a house burning.

Wrapping up

The previous sections have demonstrated the various ways events and partici-
pants can be related in causative constructions. There can be strict temporal
dependency between the cause and caused event, but this needn't always be the
case. While the periphrastic type of causative is usually more amenable to a
temporally separate reading, it can also be the case that even lexical causatives
do not have such a tight connection between the events and participants. For
example, with elemental causatives the two events are much more likely to be
intimately connected should the causer be an element (fire, cold, etc.), rather
than an animate agent making use of them.

Conversely, while periphrastic causatives will often be used in more indirect
causal situations, the example of directive causation shown in making a puppet
walk by some kind of manipulation shows that even here, the causal event and
caused event can be tightly connected temporally. (Levin, 2000) points out both
the independence of causation and telicity, and the fact that lexical causatives
cannot be reduced to a single event-type, e.g. an accomplishment.[50] The dis-
cussion and examples above should certainly reinforce this point.[51] Rather it

---

[49] This can be direct, i.e. the normal way of manipulating a puppet or via an elaborate stage
device that is manipulated to make the puppet walk. Pinocchio is, alas, an exception to this
scenario.

[50] This insight originates with (McCawley, 1976). She also discusses the importance of making
a conceptual distinction between telicity and an incremental theme.

[51] (Levin, 2000) looks only at lexical causatives and resultative constructions. The differ-
ent type of relations between the participants and periphrastic causatives are not looked at.
However, the previous sections show that the point holds in these cases just as strongly.

has been demonstrated that if one takes the various notions of causation and participants seriously a wide range of possibilities of both event-type and telicity emerge.

Rather than rely on unanalysed CAUSE and BECOME predicates to formalise causatives,[52] in the event calculus, the different types of causes and participant relations emerge as a number of different scenario types, that capture the temporal relations between the events involved, differing implications depending on how the participants are acting, as well various telicity possibilities. That different scenarios are needed not only shows that causatives also cannot be formally reduced to a single, default event-type, but also shows that numerous hybrid types also exist.

5.2.6 Formalisation

To the delight of the lazy, but attentive reader, handling the different implications requires very little modification to the standard EC apparatus. We shall first look at the levels of cause involving lexical, elemental verbs and then move on to participant control variations.

Elemental eventualities

Recall that using the proper elemental cause with an elemental verb yields something that is very close to an ordinary transitive. In fact using the dynamics for a transitive event with a dynamics and temporal duration is exactly what is needed. Take the oft-used example

(65)   Sunshine is melting the ice.

We have already seen that this type of sentence has the same implicational properties as an ordinary transitive. If *melt* is construed as telic, then the imperfective paradox holds. If atelic, there then is the implication that the sunshine did indeed melt the ice (but has not necessarily melted it away). With this in mind, it is probably best to stick to a standard dynamics for wide activities and accomplishments. What follows assumes that (65) is construed as telic in order to demonstrate all of the machinery. The scenario is quite close to that of an ordinary accomplishment, the relevant changes needed will be pointed out afterwards. See (Lambalgen & Hamm, 2005, Chapter 7), for a full explanation for a scenario of an ordinary accomplishment such as *build a house*.

First, a word about the predicate names. *Sunshine* is the warming activity of the sunshine, while $melt(x)$ is the parameterized fluent that indicates the stage of melting the ice is at. It is initially at 0 (e.g. 0 percent) at the time of the

---

[52](Levin, 2000) documents a number of author's attempts. In this paper, Levin herself is more concerned with argument linking.

scenario's beginning. This should not be taken as an absolute. For example, if the observation is made that global warming has begun melting a glacier, beginning the scenario with $melt(0)$ does not mean that 100 years a heat wave did not do the same thing. It is the current round of melting that is at issue. $Melt(c)$ is the stage at which the ice is completely melted.

The is an additional condition added of *rightpos* which means only that the ice is situated in such a way that the sun will melt it. In general for these sorts of instrumental causes this can be seen as a requirement that the object must be situated in the proper way in order to be affected by the elemental cause. For this particular scenario, *rightpos* is the most intuitive way of imagining it. This is a scenario where the object begins in a proper condition for being affected, and it is the initiation causal fluent that is started.

1. $Initially(melt(0))$

2. $Initially(rightpos)$

3. $Initiates(start, sunshine, t)$

4. $Terminates(finish, sunshine, t)$

5. $HoldsAt(sunshine, t) \land HoldsAt(melt(c), t) \rightarrow Happens(finish, t)$

6. $HoldsAt(rightpos, t) \rightarrow Releases(start, melt(0), t)$

7. $HoldsAt(melt(x), t) \land HoldsAt(rightpos, t) \land x + g(d) \leq c$
   $\rightarrow Trajectory(sunshine, t, melt(x + g(d)), d)$

To trigger the melting, the fluent *sunshine* must hold along with the added condition that the object must be situated correctly to receive its benefits. As with an accomplishments, a start event (here say the sun coming out or clouds moving away) begins the *sunshine* activity fluent and this along with the object being in the right position allows the parameterized fluent (*melt(x)*) to be released.[53]

This allows the dynamics to be active, and as long as nothing changes the ice will keep melting. Three changes can happen however, first the sun can disappear, or the ice can be moved out of the sunshine or completely melted. With the first two it may be that the ice is still melting (from a different cause) but (65) is certainly not the case and the above dynamics confirms this as the trajectory predicate would not drive the continuous change without the preconditions of both the ice being positioned properly and sunshine active.

---

[53]Alternately, the sun may already be shining and the start event is that which situates the ice properly.

If the ice has melted, the differences with an ordinary creation accomplishment such as *John is building a house*,[54] differ moderately. With the latter, once the house has reached the stage that the house is considered finished, building is terminated. Instead the elemental causative has more the features of derived accomplishment such as *John is running a mile*. It seems odd to terminate the sunshine just because the ice melted, yet obviously neither (65) or *The ice is melting* holds.

This is already taken care of and emerges from the inequality in the dynamics of statement (7). Once the ice has met the construed finish point c, the dynamic the trajectory statement in the consequent will no longer be active. As there may indeed be no ice left at the 'finish' construal point, one could add

$$Terminates(finish, rightpos, t)$$

so that it ensures that if the ice is completely melted then it trivially can no longer be in the proper position to be melted by the sun. In general the finish event of an elemental causative terminates the object condition fluent rather that the causal activity. However, this is not needed if one only wants to stop the dynamics from holding once the ice has melted.

Finally, the *start* event needs to be examined more closely. Both conditions (*sunshine* and *rightpos*) need to hold for melting to happen. For example, one of the conditions could hold initially and the *start* event is needed to trigger the other one. The sun may be out, but the event is bringing the ice outside. Here the appropriate lines would be $Initiates(start, rightpos, t)$, and $Initially(sunshine)$. The reverse is the case in the above scenario (e.g. the ice is exposed and the clouds clear), the *start* event does begin *sunshine*. However, should the scenario be such that the start event is obtaining the correct condition, (7) ensures that $melt(0)$ is still released, allowing it to be used in the trajectory predicate.

The transitive in the progressive

The progressive *Sunshine is melting the ice* patterns exactly as an activity or accomplishment (this difference depends on whether *melt* is used telically or not). Telic or not, the progressive used with an elemental agent ensures that some of the ice was melted. Should *melt* be taken to mean to melt completely, the imperfective paradox holds as for an ordinary accomplish. The major difference is that it is not enough to use the causal activity to fix the reference time, the object condition must also be fulfilled. Using the above scenario, the integrity constraint for the present progressive is

$$?HoldsAt(sunshine, now) \wedge HoldsAt(rightpos, now) \texttt{ succeeds}$$

---

[54]A destruction accomplishment would behave the same way.

This is the case no matter how the event started, i.e. initiating the object condition, initiating the elemental cause, or both. As either one of these fluents can be terminated (e.g. a cold snap, moving the ice), there is no guarantee that the telic goal will be fulfilled.

Intransitive version - tense and aspect

Suppose the sunshine has completely melted the ice – the intransitive version *The ice melted (completely)* is also the case. This sentence resembles a passive in that the subject is patient-like rather than agentive, though in this case the agent (or cause) is not accessible with a *by*-phrase. In effect, the focus is on the object and what happened to it. However, while no cause is specified or the existence of a cause mentioned, we know that something caused the result and use world knowledge to fill the cause in. The database is then required to create a scenario (modified for causality as seen above) that has some cause,[55] and a condition that that situates the ice to be affected.[56] Generally, the scenario for intransitives will have the same structure as the transitive counterpart. The main differences will be the specification of the actual cause, and the integrity constraints used.

Given this causal scenario, let *melt* be the event associated to it by the method of hierarchical planning as seen in (Lambalgen & Hamm, 2005, Chapter 7, Section 3), and the appropriate simple past tense integrity constraint, $?Happens(melt, t)$ `succeeds`, and the query

$$?HoldsAt(f_1, R), ..., HoldsAt(f_n, r)R \geq now, Happens(melt, R) \text{ fails}$$

This ensures that *The ice melted* entails that the accomplishment happened (in the past) and that the event is viewed as durative. This is necessary as this is a regular simple past tense and differs as expected from the perfect in that[57]

(66)   a. While the ice melted in the refrigerator, I did the dishes.
        b. #While the ice had melted in the refrigerator, I did the dishes.

The progressive, *The ice is melting*, goes through with similar caveats. The integrity constraint will have the parameterized fluent $(melt(x))$ coerced into an activity fluent. The database again comes up with the appropriate causal dynamics, with a generic elemental cause if need be. This is easily accomplished using the coercion methods (Lambalgen & Hamm, 2005, Chapter 11). $Melt(x)$ is a parameterized fluent and not enough on its own to drive a progressive. But, at

---

[55]If context is not supplied, something like *melter* will do, i.e. a generic elemental cause of melting.

[56]This is somewhat similar to the (Levin & Rappaport-Hovav, 1995) who suppress the external cause in S=P intransitives.

[57]For temporal *while*.

our disposal is the model of elemental causation given above, the difference being that the actual cause is a bit of an afterthought as the focus is on what is being effected exclusively. Nevertheless, we know some cause holds, which is exactly what the database is required to search for.

For example, we are given (by the sentence) $Initially(melt(0))$ and know that there is some f_1 fluent such that

$$HoldsAt(melt(x),\text{t}) \rightarrow Trajectory(f_1, t, melt(x + g(d)), d)$$

As the elemental cause is left unsaid, the database searches for something to unify with the f_1 fluent. As no specific causal fluent is given, it can be unified with a 'dummy' activity, e.g. *melter*, allowing the dynamics to be driven.[58] Aside from the perspectival difference between the intransitive and transitive versions, there is one other crucial difference.

In the transitive version, if the sunshine is blocked, the progressive no longer holds. With the intransitive, a variety of actual elemental causes can be operational during the melting event. For example while moving a block of ice from the hot sun inside to a hot room, the ice can be considered to be continuously melting. As the 'cause' for the intransitive is a blank slate, so to speak, both the sunshine and a hot room can be considered the external cause of the intransitive melting event. For the intransitive the integrity constraint (for present progressive) is

$$?HoldsAt(melter, now) \wedge HoldsAt(rightpos, now) \texttt{ succeeds}$$

that is, given that the ice is melting, we can infer that some outside factor is causing this, but we either do not know what it is, or want to focus our discourse only on the affected object.

What is developing here is an idea of a causative template for elemental causes. All elemental causative verbs have a default scenario that resembles the one above. In its barest form it only gives a dynamics that affects the object – real world details, context or use of the transitive can give rise to fluents (such as *sunshine*) that can be unified with the generic conditions. Personal agents are less direct and can be said to cause the entire causal event rather than only what is happening to the object. The template sketched above can then be utilised to capture this.

Personal Agency Cause

In section 5.2.5 we saw that if an elemental transitive causal verb is used with a personal rather than a elemental agent circumstances change considerably. While in a direct syntactic form, a personal agents directness is slightly less direct than

---

[58]Appropriate *start* event(s) must also be found.

what was seen above. In many cases, the personal agent can merely make the circumstances right for the event to happen or do something not temporally connected to it whatsoever. For example, the quote from the Arrested Development script has the main character 'burning down' the banana stand, yet no burning occurs. Here, George Michael's 'burning' is the preparation that, if successful, will cause the banana stand to burn, and perhaps, burn down.

This sort of backward viewpoint is not unique to personal agent causals, but is a well-used form known as the futurate. For example, one could say

(67)   I am flying to London tomorrow.

meaning that if all goes well (i.e. I make it to the airport, do not get detained as a terrorist suspect and sent to Egypt, board the airplane, etc.), I am flying on an airplane to London the following day. What is important is the intention, plan or inexorable chain of events that leads to a flight to London. Notice that at the time of speech, there is no implication that the subject of (67) is flying. (Lambalgen & Hamm, 2005, Chapter 8), models this sort of progressive slightly different from an ordinary present progressive.

There is no flying going on at speech time, but the use of the progressive indicates that some activity is happening at speech time. This activity can be seen a 'preparatory' fluent that is part of a 'preparation' accomplishment that, if completed, allows the 'flying' fluent to hold. Notice that (67), just as certain personal agent causal scenarios, has a double telicity about it. First, the preparation that leads to flying is telic and is easily cancellable; second, should the preparation succeed, then flying is implied, but the telic aspect of getting to London is still in the lap of the Gods.

So far this exactly parallels the 'banana stand' example, but there are some temporal differences. Examples of this sort of personal agency cause can more properly be said to be 'quasi' futurates, rather than pure futurates. For example, in the preparation scenario, one cannot change the temporal adverb from 'tomorrow' to 'now', as packing one's bags the night before can hardly be called flying. While, in the banana stand example, George Michael could say, 'I'm burning down the banana stand, now' while in the preparatory process of putting newspapers around it.[59]

'Now' can also be used once the fire is started and George Michael is basking in the sweet glow of familial revenge. In contrast, the futurate in (67) no longer holds once the actual flying commences, and the situation must be modelled by an ordinary present progressive. The 'quasi' futurate is modelled the same in both the preparatory and caused situation. These differences can be handled quite easily by changing the integrity constraint, which differentiates the tempo-

---

[59]There is a true futurate where the night before George Michael is gathering the proper materials for the next day. Here, using 'now' would be infelicitous, but 'tomorrow' is perfectly possible.

ral meaning difference between the futurate and 'quasi' futurate. In what follows I shall only give the scenario and constraints for the 'quasi' futurate, but point out the changes made from the ordinary futurate in order to make the differences clear.

The 'quasi' futurate progressive requires the presence of an activity fluent f_1, and a parameterized fluent $f_2(x)$ which are linked by a dynamics of the form

$$HoldsAt(f_2(x), t) \rightarrow Trajectory(f_1, t, f_2(x'), d),$$

and a condition for the occurrence of the culminating event e of the form

$$HoldsAt(f_1, t) \land HoldsAt(f_2(c), t) \rightarrow Happens(e, t)$$

where c is some constant. The event e triggers an activity fluent f_3 via a condition of the form $Initiates(e, f_3, t)$

The scenario to which the above statements are added should not imply that f_1 is terminated before e happens.

More concretely, taking the *John burn the house* scenario, the dynamics are in the structure of an accomplishment that, if successful, lead to the house burning down. The f_1 fluent is unified with the activity John is doing, while the parameterized fluent f_2 marks out the stages of the success of the preparation fluent. Reaching the successful stage of the preparation initiates at least one f_3 fluent. Doing this calls up an scenario associated with the elemental causative template discussed in the previous two sections.

Because of this, John's causal activities can initiate either (or both) of the activity fluents needed for the elemental causal dynamics. For *burn the house*, it is most likely that the *rightpos* fluent initially holds, as houses are usually stationary objects, meaning that John's activity would be to initiate an elemental cause.

With *John burn(up) the chair*, one can imagine that John throws a chair onto a already existing fire. Here, we would have $Initially(fire)$, and the preparatory activity initiates the *rightpos* fluent. One could even imagine a situation where John removes a firewall in order to have a spreading fire consume it.

Tense and aspect considerations

The temporal difference between personal and elemental agency is accounted for by allowing either the preparation or elemental/object fluents to set the reference

time. Thus, *I am burning the house down* can apply either during the preparation, or the time afterwards where I could be home having a celebratory drink.[60]

Furthermore, the use of 'burn' in the utterance invokes a sort of inheritance procedure in which the f_3 fluent is unified with an elemental causal fluent and calls up the entire causative dynamic scenario associated with it (and presented in the last section). This is the second telicity, meaning that even a successful accomplishment of preparation does not ensure that the banana stand is burnt down.

Recall that two fluents are needed as conditions for an object to be affected by an elemental cause. Either one of these can be triggered by an agent. If I burn a house down, then assume that $Initially(rightpos)$ but I could burn a chair by throwing it into an already existent fire. Here $Initially(fire)$ is the case and the fluent that is initiated is $rightpos$. However, in either case, the actual burning requires that the two fluents f_{3c} and f_{3o} corresponding to the elemental cause and the object condition, both hold. The preparatory event can initiate one or both (if one is not triggered it must be assumed initially).

So far the progressive looks like that for the elemental transitive and intransitive versions. But the difference with personal agency cause is the preparatory activity of the agent is enough for the present progressive to be felicitous. As it is enough that only the preparatory activity hold, the integrity constraint for the present progressive for personal agent causatives must be:

$$?HoldsAt(f_1 \vee (f_{3c} \wedge f_{3o})), now \text{ } \texttt{succeeds}$$

With this constraint the unwanted implication from *John is burning down the house* to *The house is burning down* disappears. If only f_1 holds, the house has yet to begin burning.

The simple past works as normal, it is just that the event (say, *melt*) is a hierarchically planned event of both the preparation event and the causative proper. The perfect is rather more interesting. The statement *John has burned down the house* implies that not only his preparatory activity is successful, but also that the house has, indeed, burned down. This means that the perfect must be modified to use two consequent states to set the reference time. Let $prep(p)$ be the consequent state of the successful preparation event and $burn(c)$ be the state that corresponds to the burned down house. The integrity constraint for the above present perfect would then be

---

[60]Further indirectness, such as *The Mafia is burning the house down* would also hold if the house is burning and the Mafia's causal role was to hire an arsonist. Interestingly, the 'quasi' futurate interpretation of this is a bit more difficulty with this scenario and it is the arsonist doing the preparation. The Mafia's preparation is one level back, and up until the house is burning, the best non-futurate expression is a periphrastic such as *The Mafia is having the house burned down.*

$?HoldsAt(prep(p) \wedge burn(c)), now$ `succeeds`

That is, both consequents must hold to account for the above. It is not enough that the house merely burns down, but it must be seen as caused by something John did. This may initially seem a bit odd, but it reveals the nature of control between the participants that has already been seen.

Participants and Control

In the situations looked at in the last section, the objects such as houses, banana stands and chairs have very little control over what happens to them. With more animate participants this may not be the case, as seen in (61). For convenience this set of examples have already been seen in (21):

(68)   a. I've walked Mary to school.
        b. I('ve) had/made Mary walk to school.
        c. I ('ve) made the puppet walk.
        d. I ('ve) made the puppet walk across the stage.

Taking the most obvious situations corresponding to (68-b) and (d). In the former, my causal job is to convince or order Mary to walk to school, what happens afterwards is beyond my remit. In the latter, the inanimate puppet has very little choice in his movement. No matter how directly or indirectly I cause the puppet to walk, the only participant with a degree of control is me.

The scenario where the causee (Mary in this case) has control over her actions is quite similar to that of the personal agency causatives for elemental eventualities. In both cases there is a degree of indirectness. But, as the implications for the two types of constructions are different involving the control relations between the participants, there will be some vital differences.

The scenario consists of a preparatory accomplishment on the part of the causer (the *making*), and the eventuality caused corresponds to the complement. The caused event is, of course, quite different, in (68-b) being an ordinary accomplishment. Furthermore, the causal event needn't cause any action on Mary's part whatsoever, only the assent that she will walk to school. Parallelling the indirectness of the 'quasi' futurate, the causal event will be an accomplishment that triggers an activity fluent. Creatively, the fluents are called *make* and *made* to stand for the preparatory fluent and the stages of a hopefully successful making event.

Sentences of the form (68-b) to (68-d) require the presence of a preparatory event that which are linked by a dynamics of the form

$$HoldsAt(made(x), t) \rightarrow Trajectory(make, t, make(x + g(d)), d),$$

and a condition for the occurrence of the culminating event e of the form

$$HoldsAt(make, t) \wedge HoldsAt(made(c), t) \rightarrow Happens(e, t)$$

where c is some constant.   The event e triggers an activity fluent f_3 via a condition of the form Initiates(e, f_3, t)

Here, the f_3 fluent will be *walk*, with the eventuality for *walktoschool* inherited. With the more complicated version, where Mary merely assents to walking to school, the f_3 fluent would be a preparation fluent for Mary, with the eventuality being a futurate version of *walktoschool*. This basic scenario is useful for types of causation both where the causee still exercises a degree of control and those where she does not. The difference is modelled in the integrity constraints for the two types of situations

The Progressive

In both control situations, the progressive *I'm making Mary walk to school* needn't imply that Mary is walking now. The difference being that once Mary starts walking, the activity of the no-control version stops, while the control version (dragging Mary for example), must continue lest Mary's activity stops.[61] In both cases the reference time for the present progressive is set as follows:

$$?HoldsAt((make \vee walk), now) \; \texttt{succeeds}$$

that is, as the causer, *I'm making Mary walk to school* is felicitous if I'm in the midst of the preparatory activity or Mary is actually on the way to school. In general, the control differences between the participants can be illuminated by the addition of one scenario statement each for the two different situations. For the no-control version.

$$Terminates(e, make)$$

For the control version

$$?\neg HoldsAt(make, t), HoldsAt(walk, t) \; \texttt{fails}$$

Intuitively, the fluent *make* is quite different in the two cases. In the first case, *make* is most likely an act of convincing or ordering Mary to walk to school, i.e. manipulative causation. The second case, where I physically force Mary to walk, is an example of direct causation. Here, should I stop my causal activity, *I'm making Mary walk to school* no longer is felicitous.

---

[61]Furthermore, the control version doesn't allow for assent on the part of the causee, naturally.

Simple past and perfect

This distinction of whether the causee has a degree of control or not can be seen with both the simple past and the perfect. Sentence (68-b) only implies that the preparation event was a success and implies nothing about whether or not Mary walked to school. Sentence (68-d) does imply that the puppet walked across the stage. The agent participants contribution can be either as a quasi-futurate, where I push a button that starts a walking mechanism or something more direct, where I manipulate the puppet in the normal fashion.[62]

Contrasting (b) and (d) in their more indirect scenarios, the formal differences of participant control can easily be seen. Both scenarios will have a preparation event and the caused event associated with the complement. These two subevents will each have a consequent state associated with a successful completion. For situations such as associated with (68-d), the implications of the perfect are the same as with the *burning* examples seen in the last section. Formally, they are also the same. The simple past goes through with no modification, and the perfect, as earlier, requires two consequent states to hold. Let $make(c)$ correspond to the consequent of the preparatory event, and $walk(c)$ to correspond to the state of the puppet having walked across the stage. The reference time for the present perfect version of (68-d) would then be

$$?HoldsAt(make(m) \wedge walk(c), now) \texttt{ succeeds}$$

In general, in cases where the causee has no control, the perfect requires that both the state that corresponds to the causer and the causee hold. In a sense, the causee participant is irrelevant for what happens and the consequent must be appended to what the causer does.

(68-b) in both simple past and perfect forms has implications only for what the causer does. This is a bit more straightforward, in that it is only the tensed verb that is involved in setting the reference point, the complement being under no obligation to happen. Thus the implications of the both perfect and the past only go as far as the causer's involvement. For the perfect, the 'I've made' in (d) only implies that the preparation was successful. This is easily expressed as

$$?HoldsAt(make(c), now) \texttt{ succeeds}$$

The simple past is actually simpler than those where the causee has no control. Sticking with the intuition that in these situations the past tense, 'made' only

---

[62]The second would ignore a preparation fluent and insert my activity directly into the dynamics. This has the pleasant consequence that the *making* and *walking* need to be temporally co-extensive, causing the implications of the puppets successful walk to come for free. This situation could also be felicitously expressed with the lexical 'I walked the puppet across the stage'.

implies the success of the preparatory event, it is indeed only the scenario for this that is turned into a hierarchically planned event for the appropriate integrity constraint.

For true[63] progressive aspect, both types of participant situation work exactly like the burning examples in the last sections. The requirement on the reference time is that either the preparation or caused fluent hold at the reference time.

States

States, as we have seen, can also be caused (as often happens in resultative constructions), but another more interesting phenomenon often happens, where the basic causative construction is utilised, not with a 'real' (i.e. conventional) verb, but with a stative adjective. Typically, one would express state causation using a periphrastic formulation, e.g.

(69)   John made me nervous.

Of course there is no acknowledged verbal analogue to *nervous*, however I have begun noticing a certain usage. Apparently among some young British people, *Nervous* is a verb, whose use seemed to popularise when a Big Brother contestant was fond of it. Oddly, it has both a transitive and intransitive use (which is not possible with *scare* or *frighten*). So both *You're nervousing me* and *I'm nervousing* (as in getting or being nervous) are somewhat common among a certain linguistic group. My dialect has a lovely verb *unnerve* (only as a transitive) which blocks this awkward verbal form for me. Personally, I find *nervousing* rather ugly, but there is a chance it could catch on in a broader population.

Final Remarks

While the above three sections do give quite a number of examples of different types of causal situations and their accompanying causative expression, this is by no means meant to be an exhaustive survey. The sections on the semantics of causation and the cross-linguistic excursions should be enough to show that the various combinations of direct/indirect, directive/manipulative and control between the participants is legion.

What has been done is to provide the machinery and general principles behind handling the various forms imaginable. Should the reader come up with a causal situation that has not been handled above, they can amuse themselves by using the above methods to formalise it for themselves. *Break*, for example, is a vague enough causal to allow an incredible amount of situations in both its direct lexical and periphrastic uses.

---

[63]That is, real past or present progressive - excluding futurates and 'quasi' futurates.

In a general sense, there is an aspect of these causatives has yet to be formalised. The manipulative/directive distinction fell out in the integrity constraints, but the is a (syntactically) formal difference that I can't account for. With the *John burned the house down* scenario, either the direct form or *John made the house burn down* would be formalised the same way.

In one sense, this is exactly what we want. One can use different forms for the same situation depending upon both intentions of the participants and they way the speaker wants to convey the situation. Therefore, the formalisation should, for the most part, be the same for both forms.

But, there are a few differences. Where there is intention one would more likely use the direct form if everything goes to plan in a normal way. But, even where there is intention, but the causer had to do more than he expected to burn down the house, the periphrastic would more likely be used. Similarly, a child who plays with matches and causes a calamity could have their behaviour described in either way. The direct form is much stronger rhetorically. These differences would need to be elicited in a broad pragmatic theory. While with some verbs, a direct/indirect casual distinction may manifest itself iconically in a mapping to, say, a lexical/periphrastic causative distinction.

Speculatively, I would suggest that once there is a distinction between direct and indirect causation in the language, it could be the case that the expression of directness is used not only for the spatio-temporal situation, but for implications of the causer's responsibility, intention, in contrast to indirectness used for accidental causation and the like. As we shall see, sometimes the distinction disappears (over time) altogether, and a direct, lexical form can be used for any type of cause whatsoever.

As a coda, a brief diversion follows that examines a few verbs that are paragons of verbs that are considered externally caused and involve elemental causation. The range of causes applicable increase in both type and temporally proximity in a consistent way over the course of time.

5.3   Coda: Diachronic development of causal level

This is a small unscientific survey into the history of a few causatives that attempts to support the idea sketched above. This is no corpus study, but an excursion into the Oxford English Dictionary (*Oxford English Dictionary*, 1989) looking at the development of a few causatives and their expanding acceptance of both types and level of cause. To begin, it may be useful to begin with a quote from the OED itself as to the change in *kill*

> To put to death; to deprive of life; to slay, slaughter. In early use implying personal agency and the use of a weapon; later, extended to any means or cause which puts an end to life, as an accident, overwork, grief, drink, a disease, etc. (p.427)

The cause expands to encompass anything, as long as the causative meaning is maintained. This allows non-animate causes, indirect causes, or event ideas, feelings, and various activities. The early restrictions required personal agency, in a rather direct way, to be appropriate. If we recall the difference between the subtle distinction of direct causation between John or fire burning down a house something interesting happens historically. The transitive version of *freeze* has its first examples of a direct, elemental cause, and only much later shows examples of animate subjects using such causes.

Freeze (p.170): "Of natural agencies: To change (a fluid) to a solid form by the action of cold...Also said causatively of personal agents."

The earliest example of this use is from 1494 (Fabyan, *Chron.* VII)

(70)   In this...yere...began a froste that..frose the Thames.

Only in 1781 is there an example of personal agency. (Cavallo, *Phil. Trans.* LXXI 516 I)

(71)   I have froze a quantity of water with an equal weight of good ether.

This pattern occurs also with *age*, and I would presume many causals verb where the causee is subject to effect by natural agencies.[64] It is standard to envision an agent directly using natural means (or created means) as directly causative. Yet with the above verbs, this distinction may indeed of been made early in the verbs life.[65]   Synchronically, one absentmindedly leaving the refrigerator door open would be cause enough for an exasperated flatmate to exclaim "You melted the ice cream." As we have already seen, this subtle distinction has major implicational differences when examining these sort of causatives in the progressive aspect. Acknowledging this is essential to get the formalisation correct.

Finally, the verb *disappear*[66] is traditionally thought to not have a transitive use. In fact, the sentence

(72)   *The magician disappeared the rabbit.

seems to be quite popular in testing the acquisition and over-generalisation of argument structure.[67] However, sentence (72) is not an over-generalisation that

---

[64]*Melt* and *Burn* originally had slightly differing verbs for transitive and intransitive uses based on strong and week verb forms. Eventually they merged into a single form and became prototypical English lexical causatives. However, it is unclear from their examples as to whether there were any causal level and participant restrictions, as seen in *kill* and *freeze*.

[65]*kill* as part of a suppletive pair as well as initially having a more specific semantics is slightly different.

[66]This is one I find to be difficult to determine whether it can be considered externally or internally caused, or even S=A or S=P.

[67]See, e.g. (Tomasello & Brooks, 1999). This is paper is but one of many that use this example to see when and how children make 'errors'. (Levin & Rappaport-Hovav, 1995) give a detailed account of verbs of appearance and disappearance and conclude that they are a rather

children make, but is even attested in the OED. They have an entry for a transitive use, but the first instance is relatively recent and far later than the entry of the verb into the language.[68]

(73)   a. We progressively disappear the faces of the dodecahedron. (1897)
        b. The magician may speak of disappearing or vanishing a card. (1949)[69]

In current times it is certainly not uncommon to hear talk of malevolent governments disappearing political prisoners.[70] Moving into the realm of bad puns, the technology website (www.theregister.co.uk) entitled an article about an program that completely erases incriminating e-mails,[71] *Honey, I disappeared the e-mails.*

What these examples mean for theories of argument structure acquisition, or the semantics of verbs of appearance remains to be seen.

5.4   Resultatives

This brief section on the class of Resultatives is in no way meant to be an exhaustive exploration of the subject.[72] Rather, we shall look at the basics, before examining their interactions with lexical causatives, implications rising from the complexity of events, and the causation of the construction in general. The main point of the next subjection is that simple compositionality cannot account for both the syntactic oddities and semantic coercion that shows itself in the resultative construction. The clever argumentation in what follows is from one of the pioneering works arguing for a constructional approach to language, (Goldberg, 1995).

5.4.1   Basics

We will begin with some examples of the most common resultative types.

(74)   a. John hammered the metal flat.
        b. Jane sang the baby asleep.
        c. Sally cried herself into hysteria.
        d. Think yourself thin!

---

special case in that the notions of external or internal causation play no role. However, they are said to have no transitive, which while not the case for *disappear* does seem to hold for *appear*.

[68]1530 is the date of the first example of intransitive *disappear.*

[69]This is from the journal, American Speech!

[70]In fact, I thought this is where the usage came from before happening upon the OED entries. The poor souls are often called the 'disappeared', which, to my mind, made the transitive use fairly natural.

[71]A boon for individuals and corporations who have to respond to a discovery summons.

[72]For this, the reader is recommended to see various approaches in (Goldberg, 1995, Chapters 7 and 8), (Rothstein, 2004, Chapters 3 and 4) and (Levin & Rappaport-Hovav, 1995).

Sentence (74-a) can be considered the most prototypical, stating that the result of John hammering the metal rendered it flat. On the surface, (74-b) is rather similar in meaning (John's singing caused the baby to sleep), except that *sing* is normally an intransitive verb.[73] Thus

(75)   *John sang the baby.

(74-c) is rather similar in nature, but is an example of a reflexive resultative. Both (b) and (c) can be called 'fake object' resultatives. (74-d) is an example of a sort of imperative/fake object resultative construction.[74]

Resultative constructions also exist in intransitive forms, though there are restrictions to their applicability.

(76)   a. The river froze solid.
        b. *The river froze itself solid.[75]
        c. John sang himself hoarse.
        d. *John sang hoarse.

S=P type verbs can only be used in an intransitive resultative construction, while S=A intransitives, need an object to participate in the resultative construction. We shall see the implications of this in the next section, but shall first examine Goldberg's more general claims about the construction.

First of all, she examines the claim that the post-verbal NP in the 'fake object' constructions is not a real argument of the verb.[76] The claim against true object status of the 'fake' objects comes from a series of tests comparing standard sub-categorised objects and fake objects in regard to the middle formation, adjectival passive and nominalisation, which are considered evidence for argument-hood.

(77)   Transitive Resultative: He hammered the metal (flat).
        a. Middle Formation: This metal hammers flat easily.
        b. Adjectival Passive: the hammered-flat metal.
        c. Nominalisation: the hammering of the metal flat.

(78)   Fake Object Resultative: He drove his tires *(bald).
        b. Middle Formation: *Those tires drive bald easily.
        c. Adjectival Passive: *the driven-bald tires.

---

[73]When it is sub-categorised for a (syntactic) direct object, the object is not a human, but a song-like object.

[74]'Fake object' resultatives also passivise, as in *The tools were wiped clean.*

[75]The differences between (a) and (b) are not confined to the resultative construction, but more to there being an 'S=P' verb. Removing the resultative XP from the sentence, **The river froze itself* is reportedly as bad. In both cases, without help from context, it is rather difficult to conceive the river as causally efficacious for freezing. However, if it is, the (a myth or science fiction context) ungrammaticality disappears. Later in the section we shall see how hyperbole and metaphor is enough to 'agentivise' patients and make them felicitous in these constructions.

[76]All examples taken from (Goldberg, 1995, p. 182 - 185).

        d. Nominalisation: the driving of the tires bald.

However, she notes that the paradigm in (77) does not apply across the board to all transitive verbs with true objects.

(79)   a. Middle Formation: *Pat kicks black and blue easily.
           b. *The washer loads full easily.
           c. Adjectival Passive: *The washed-shiny-clean face.
           d. *The shot-dead man.
           e. Nominalisation: *The washing of the face shiny clean.
           f. *The driving of him crazy.

Thus, while felicitous distribution in these constructions (as in (76)) may be evidence for being an argument, the converse is certainly not the case, as *washer* is a (subcategorised) object of *load*, *Pat* is the object of *kick*, etc. If there is no reason to consider the objects in (79) as 'fake objects' (despite them failing the tests), then there is no reason to consider those in (78) to be anything but real objects.[77]

    Finally, Goldberg notes that context has much to do with their acceptability, and has examples where 'fake object' resultatives can participate in the above constructions, the middle construction, at least. For example, in a situation where someone in charge of props for a movie need to drive 50 tires bald for a stunt, the following suddenly becomes acceptable.

(80)   Go buy some cheap tires for that scene, those inexpensive tires drive bald really quickly.

Once this is acknowledged, the constructional approach becomes clear. A simple, lexical compositional approach will not do. For example, one would need to have special lexical rules for the fake-object resultatives. But, as the possibilities of productivity are so varied, endless new senses would need to be postulated for intransitive verbs, or transitive verbs that take objects in the resultative constructions that they normally do not, as in *They drove the tires bald.* Moreover, the basic sense of verbs can also undergo a drastic change as with

(81)   Shelly read *Jane Eyre* to tatters.[78]

which is about the physical effect Shelly's activity has on a book, whereas in *Shelly read* Jane Eyre the book is not such a force recipient. Given this, Goldberg posits two resultative constructions which have a meaning (for the transitive) that an agent does some activity to an object that changes its state. The result XP (in this case 'to tatters'), specifies the result, and with world knowledge, gives

---

[77]Nevertheless, we will continue to call them 'fake object resultatives' out of convenience.
[78]Example from (Rappaport-Hovav & Levin, 2001).

a clue as to how the activity effects the object.[79] The intransitive resultative construction (*The river froze solid*) has only a patient, that is the object of a change-of-state verb, with the result XP specifying the change-of-state in more detail.

5.4.2 Event Structure Implications

Transitive vs. Intransitive Resultatives

Recall the difference in the following

(82)　　a. The river froze solid.
　　　　b. *The river froze itself solid.

Here verbs that undergo the causative alternation (as S=P types) can only appear in resultatives in the intransitive construction. Whereas, verbs that are typical S=A intransitives can only appear in the transitive construction. Sentence (82-b) can be seen to ungrammatical in an intuitive way, first of all. We have already seen that the object in the transitive construction is the patient that is acted upon, ultimately ending up in the result state. However, rivers are generally recipients of freezing and not the instigators. Looked at in an event structure way, it is also the case that there simply are not enough event slots, if we maintain that intransitive *freeze* is an S=P verb. In (65) we saw that the participant *ice* (for *The ice melted*) was linked to a parameterized f_2 fluent,[80] with a dynamics associated with an abstract cause. So, if an abstract cause is associated with the f_1 fluent, the (semantic) object of the verb as an f_2 fluent, there is nothing left to link the 'itself' to.[81] However, it is certainly possible to have such elemental causatives appear in transitive resultatives, the condition being that the grammatical subject must be construed as an agent (effecting a cause at least), and the syntactic object is a patient.

(83)　　a. The blizzard froze the river solid.
　　　　b. The air conditioner froze itself solid.

The only problem then with the ungrammaticality of (82-b), then is that the river is not conceived as causally efficacious (as far as freezing goes). If it could be so

---

[79]This is the Direct Object restriction, which states that it is the object that undergoes a change of state in resultatives. As analysed in (Rappaport-Hovav & Levin, 2001), there do exist subject-oriented resultatives (*They followed the star out of Bethlehem*). We shall ignore these, along with sentences of the form, *John drove the car drunk*, which are not resultative at all, but merely adjectival predications of the subject at the time of the event . For an analysis of these, see (Rothstein, 2004).

[80]*The river froze* is identical in structure.

[81]We take (transitive) resultatives to be basically accomplishments, leaving us with only the options of a punctual event and result state fluent to link 'itself' to. Neither option makes a bit of sense.

conceived, then it should be as felicitous as (83-b). Indeed, this is exactly what happens (from an example taken from Google):[82]

(84)   A bottle of country club soda froze itself and exploded in my fridge tonight in a violent act of protest.

A bottle of soda is probably as capable of being causally freezing as a river, yet with the right context, can sound perfectly felicitous. Note that the speaker, metaphorically notes that the soda was making an act of protest, something normally confined to agentive beings.

The contrast in

(85)   a. John sang himself hoarse.
        b. *John sang hoarse.

sheds light on event structure in a slightly different way. *Sang* typically carries no hint of force transmission, but the result XP, *hoarse* implies the existence of a 'becoming hoarse' event.[83] As *sing* is typically associated with an f_1 fluent, the link to a canonical e (to give the change of state) requires the existence of a dynamics that includes an f_2 fluent as well. This subevent then requires an argument XP in the syntax to correspond to the parameterized fluent, which is exactly the purpose of the 'himself' in the reflexive resultative in the grammatical (85-a). Thus, we then end up (after the coercion) with a garden-variety accomplishment, where John's singing over time, eventually makes him hoarse.[84]

There is *The gate rumbled shut* which is durative, has a change of state, but no object. Here is an example of how many motion verbs with inanimate subjects can be conceived as either externally caused or the cause can be ignored and the inanimate object seen as the source of movement.[85] In this case the former view is taken, and as with *The ice melted into a puddle*, the dynamics[86] is already given by the verb, with the result XP contributing the telic endpoint.

---

[82]Note that there is no result XP in the example. But if the reader should accept it, adding a result state like *solid* changes nothing. Also, this point is not specific to resultatives, but these sorts of causatives in general.

[83]We are following in spirit the Argument-Per-Subevent Condition as stated in (Levin & Rappaport-Hovav, 2004) which states that 'There must be at least one argument XP in the syntax per subevent in the event structure. In event calculus terms, f_1, f_2 and canonical e's are considered subevents. The possible exception is when dealing with intransitive S=P verbs such as 'The ice melted.' We postulate the existence of an f_1 linked dynamics, but this sub-event is an unrealised, abstract cause, and should be seen slightly differently.

[84]This needn't be only the case for reflexive resultatives, but works the same for 'John sang the baby asleep', though the participant corresponding with the f_2 fluent is now a baby whose wakefulness is changing.

[85]The latter view is taken with such examples as in (10) and (13).

[86]The scenario statements involving both f_1 and f_2 fluents and their interactions. If *roll* or *rumble* are conceived as self-contained with inanimates, their modelling will be the same as for (85).

Finally, (Rappaport-Hovav & Levin, 2001) note that certain verbs can appear in either the transitive or intransitive resultative construction, but with different temporal implications.

(86)   a. She kicked free.
         b. She kicked herself free.

The result is the same, but the temporal profiles are different. In sentence (86-a), the situation is instantaneous, while the in (86-b), the situation is understood to take some time. Happily, this emerges effortlessly in the event calculus. In both sentences, the result XP implies that there is a 'become free' event, which requires a canonical e, by itself, or a dynamics that results in such an e. The first sentence can be modelled as an achievement, which consists of a punctual event *kick* and the resulting state, *free*. As achievements are instantaneous, the temporal profile of (86-a) follows directly. Unlike, *sang* in *He sang horse*, *kick* is interpretable as a instantaneous event (a single kick), which is exactly what we need to get the achievement-like structure).

Sentence (86-b) contains two arguments, leading us to model it along the lines of *He yelled himself hoarse.* The presence of the reflexive leads us to postulate a dynamics that leads to an event of becoming free. As kick can be construed as an iterative activity, it is this that drives the dynamics that drives the process of becoming free. As dynamics take time (as opposed to canonical e's), the event as a whole takes time.

Further complications

In (Rappaport-Hovav & Levin, 2001), a number of ungrammatical non-subcategorised transitive resultatives that are ruled out due to an already saturated event structure.

(87)   a. *The bomb exploded the watermelons into the air.
         b. *The ice melted the floor clean.
         c. *The water evaporated the pot dry.

We shall pay the most attention to the first example as it is so cleverly constructed. Their explanation is that since the above verbs are externally caused (S=P), which means that while there is no expressed cause, it nevertheless uses the same event structure. The addition of the result (watermelons into the air) requires the addition of another subevent that is not available. But this is where things get interesting. The ostensibly similar

(88)   The bomb exploded the watermelons into pieces.

is perfectly grammatical.[87] We can see the difference by teasing out the implications. Certainly an implication of (88) is that

(89)   The watermelons exploded.

meaning that *exploded* can be construed as either externally or internally caused. In the latter it is a simple, internally caused S=A verb, that can have causal efficacy such as having a bomb explode watermelons. (88-b) can be modelled along the lines of a simple resultative, meaning that the event structure is not over-saturated.

But if *explode* were considered to be an S=A verb in (87)(a), it would also have the implication that the watermelons exploded. But this is not the case. As far as I can tell, the result XP 'into the air' is more associated with the watermelons flying whole and undamaged into the air. In other words, the bomb explodes, and the watermelons fly into the air from the blast of air created by the explosion. This is too much structure to put into a single resultative, even when explodes functions as a transitive. The latter two examples of (87) are clearly internally caused.[88]

5.4.3   Causality

While the resultative structures we have been look at in this section are considered to have causality as part of their meaning, (Rothstein, 2004) begs to differs slightly. While she agrees that a sentence like *John sang the baby asleep* is interpreted causatively, it is more of a pragmatic inference than part of the semantics. Her natural language version of the formalism goes as follows

> The singing event e_1 was assigned a telic point e_1' and the telic point of the singing event was time-participant connected to the event e_2 of the baby being asleep.

That is, the semantics only says that the singing coincides in a certain way with the baby going to sleep.[89] The causal 'meaning' is then inferred in a sort of Humean way from the coincidence. While this well may the way we infer causality in situations, this is not to say that the construction we use shows whether we view a given set of situations as causal or not. To explore this point, we will look at a number of examples from Rothstein that she says are resultative, but are to be interpreted non-causally.

---

[87]This is not addressed by (Rappaport-Hovav & Levin, 2001).

[88]It takes much more context to have *melt* or *explode* be S=A intransitives, and they are clearly not in these examples.

[89]It actually needn't be the case that there is such a tight temporal connexion between the activity and the result achieved. In their exhaustive typology of resultatives (Rappaport-Hovav & Levin, 2001, p.775) note examples such as *Sam sang enthusiastically during the class play. He woke up hoarse the next day and said, 'Well, I guess I've sung myself hoarse.'*

(90)   a. The crowd cheered the gates open.
         b. Every night the neighbour's dog barks me asleep.
         c. On May 5, 1945, the people of Amsterdam danced the Canadians to
         Dam Square.
         d. Mary drank John under the table.

For (90-a), she allows that the situation could be causal if there was some effect
of the shouting directly upon the gate (a voice activated lock, e.g.). In other
circumstance it is just a situation of the crowd shout and the gates opening, which
we may take as causal, but really isn't. However, we have seen that such a direct
notion of causality is but one of the many ways cause can happen. For example,
a journalist would use (90-a) in a situation of political revolution, where a crowds
demands effects the guards so much that they change their political allegiance
and open the gate to the palace. Such a situation would still be considered causal
in terms that were explored at the beginning of this chapter.

   But more importantly, using the resultative construction rather than a differ-
ent paraphrase does express some form of causality. Imagine the scenario above,
but this time the gates are always set to open at 10:00 a.m. But the crowd is
also screaming at 10:00 a.m., which could give one the impression of causality. If
the journalists doesn't know this fact, he may indeed use (90-a), in order convey
people power. But examine the incongruence if he does know the truth.

(91)   a. #The crowd cheered the gates open, but it was only a coincidence.
         The gates always open at that time.
         b. The crowd appeared to cheer the gates open, but it was only a coin-
         cidence. The gates always open at that time.

This discourse makes little sense to me – explicitly denying the causal link renders
the first sentence a bit odd. In contrast, had he merely said tha the crowd
cheered and the gates opened, but it was only a coincidence, he would be denying
a possible inference one could make, but is not expressly part of the meaning.
Causality can also be denied within the construction as in (91-b), by explicitly
doubting the resultative.

   Sentence (90-b) is not causal, but it is ironic, and what is ironic is that the
dog's barking does not stop the man from falling asleep. The irony only works
because the ordinary use of such a resultative is causal. Sentences (91-b) and (c)
are still causal, but, again, most likely not directly causal. It is doubtful that the
people of Amsterdam physically forced the Canadians to dance, but using this
construction conveys that the gratitude and enthusiasm of the newly liberated
peoples was the cause of the mutual celebration.[90] Similarly, while one does not
need to have alcohol poured down one's throat in order to be drunk under the
table, the expression does put a bit of responsibility (or blame) on the part of

-----
[90]Compare with the blander, *The Dutch and Canadians danced together to Dam Square.*

the subject. The causality, while not direct, can be seen as someone trying to keep up with another's drinking and failing. The causality is perhaps the most metaphorical here, but would belong to the realm of social causality.[91]

However, this is not to say that all types of resultatives are necessarily causal. It may be the case that ordinary 'change of location' resultatives may not be causal. Rothstein gives the example
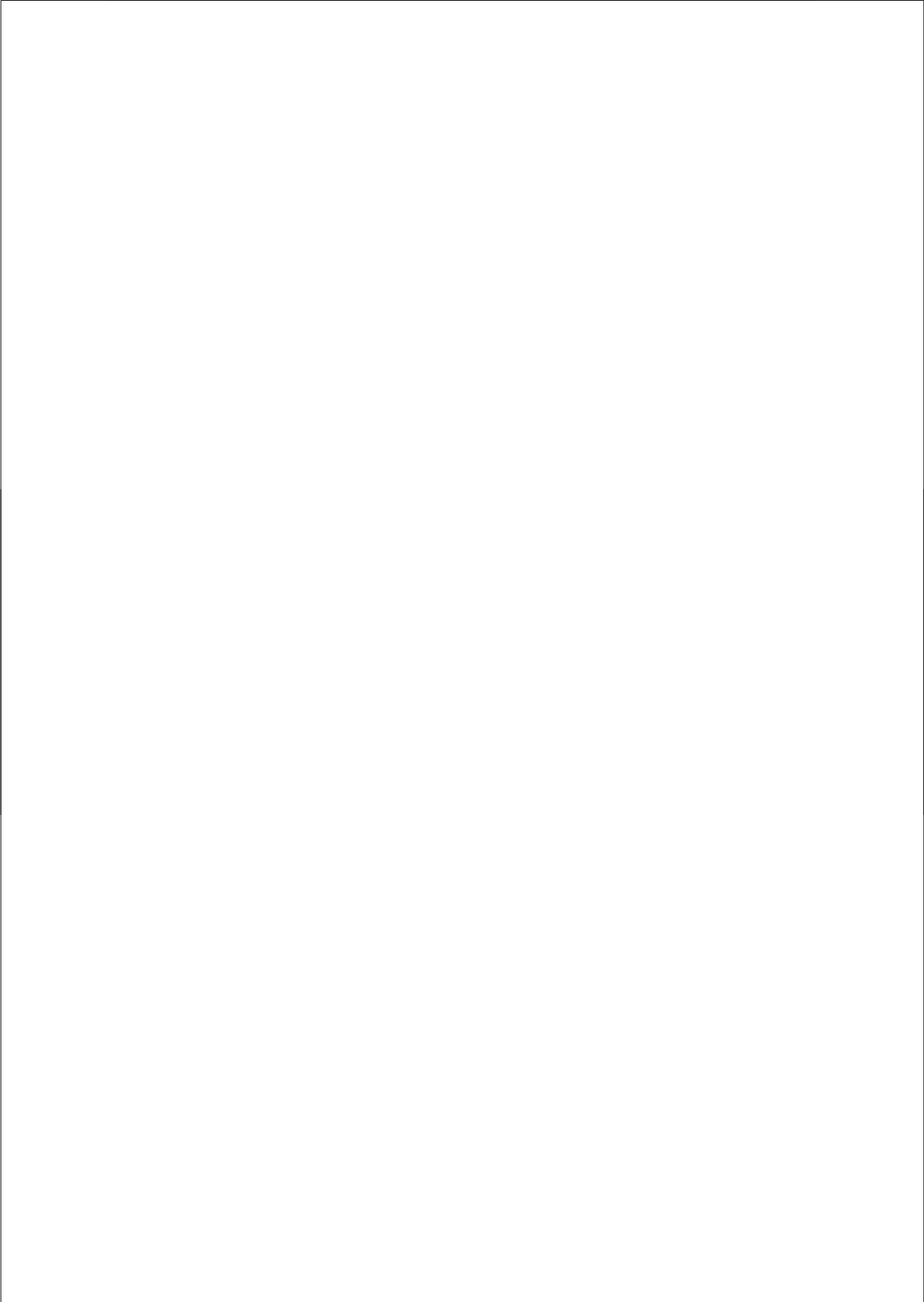
(92)   The crowd applauded him off the stage.

which, while having a causal interpretation[92], can also be interpreted as simply meaning that while he left the stage, the crowd applauded. I have no explanation why this type of resultative should not have as core a causal meaning as the others, but the culprit cannot be said to lie with the choice of verb as the following is certainly causative.

(93)   The crowd applauded him into glowing reviews.

---

[91] I have been in this exact situation. Not wanting to be out-drunk by a woman, I would stupidly try to keep up the pace. The next day, in my agony, she would always get the blame. In general this idiom is used not merely to imply that two people drank alot together, but that there was some kind of contest or competition going on.

[92] This is easier to see with *The crowd booed him of the stage.*

Chapter 6

Nominalization

Nominalisation is, in general, the process of turning verb phrases into eventuality denoting nouns. While the previous chapters involved using event-types and fluents for formalisation of single sentences or multi-sentence discourses, with nominalisation, these entities are denoted by a special type of noun phrase, which can then be embedded in a larger structure.

This chapter will specifically look at (*Ing*) nominalisation in English, beginning with the received wisdom about the basic facts of their syntax and semantics. An excursion into the past is then made, examining its history and development. With that background in mind, a closer look at the received wisdom is made, in both the areas of the nominalisation's syntax and categorical status, as well as a broadened view of their semantics and interpretation. The chapter will conclude with some of the more 'novel' readings formalised in the Event Calculus.

6.1 The Standard Story

6.1.1 Syntax

(Abney, 1987) distinguishes between four different classes of gerunds:[1]

a) Acc-ing: John singing the aria.
b) PRO-ing: singing the aria
c) Poss-ing: John's singing the aria.[2]
d) Ing-of: John's singing of the aria.

---

[1]It should be noted that there are numerous other uses of V + ing, including the progressive (*John is crossing the street*), participle (*John, being late for work, decided to take a cab*) and as a nominal modifier (*That was a shocking film*). The first two of these constructions will come back to haunt later.

[2]This is not meant to imply that the possessive in Poss-ing gerunds is a 'real genitive' like *John* in *John's car*.

The origins of the names are fairly obvious given that the subject of the gerund receives accusative case in Acc-ing and possessive in Poss-ing . PRO-ing, as the gerund has no overt subject and receives obligatory control, as in *John$_1$ enjoyed PRO$_1$ reading the book.* The name *Ing-of* seems to speak for itself.

First off, we shall see the distributional properties that all four have in common – a distribution that is characteristic of NP's i.e. any distributional slot that an NP can appear in, a gerund can appear as well. All four gerunds occur as subjects, direct objects, and prepositional objects.[3]

As subject:

(1)   a. *Your having broken the record* was a surprise.
        b. *Them trying to sing a song* was just too horrible.
        c. *Singing arias properly* is difficult.
        d. *John's singing of the aria* is not to be missed.

As object:

(2)   a. The hunchback hated *a nice lady being hanged.*
        b. John enjoyed *reading the book.*
        c. I disregarded *this insulting his opponent.*
        d. We deplore *the killing of innocents.*

As object of preposition:

(3)   a. Michael counted on *them finishing the book soon.*
        b. Mary was obsessively afraid of *being ill.*
        c. They didn't approve of *my leaving without a word.*
        d. The football match ended with *the sending off of two players.*

As far as their syntactic distribution goes, all four gerund types behave more or less the same;[4] however, types (a)-(c) do differ rather dramatically to (d) in

---

[3]Abney also gives other distribution positions such as topic position, cleft position, and subject of a sentence following a sentence-initial adverb such as *perhaps*; however the above examples should suffice to illustrate the "NP-like" qualities of the various gerunds. Note that for and that clauses cannot appear as objects of preposition nor any of the other positions just mentioned. From Abney:

(i)   a. I learned about John('s) smoking stogies. (Acc-ing and Poss-ing)
        b. *I learned about that John smokes stogies. (that-clause)
        c. *I learned about for John to smoke stogies. (for-clause).

.

[4]Distribution is not the only feature of external syntax. Variance also occurs among the four gerunds in regards to agreement, pied-piping and scope. In these cases, Poss-ing and Acc-ing part company. Also note that there is a small class of (semantic) verbal contexts in which only (d) types occur – this is discussed in section 1.2.

regards to internal syntax, the former possessing internal syntactic properties of a verb phrase, and the latter possessing those of a noun phrase.[5] Therefore, when convenient, I will refer to Acc-ing, Poss-ing and PRO-ing as verbal gerunds, and Ing-of as the nominal gerund.

Verbal gerunds (if transitive) take a bare NP object, do not allow determiners or quantifiers, can be modified by *not*, allow auxiliaries for tense and voice, are modified by adverbs (not adjectives), and in the case of PRO-ing, a PRO subject is obligatory.

(4)   a. John's finishing the book was surprising.
       b. *The/Every/No reading the book was amusing.
       c. Not reading the manual properly was the reason John lost a finger.
       d. We were all thankful for John's having left the room.
       e. John was unhappy with being so hated.
       f. John's singing the song loudly offended many an eardrum.
       g. *John's loud singing the song...
       h. John enjoyed PRO reading the book.

On the other hand, Ing-of gerunds take PP complements (never bare direct objects), take any determiner or quantifier, do not take auxiliaries for tense and voice, take preposed adjectives (not adverbs), and do not have a PRO subject.

(5)   a. John's reading of the book was splendid.
       b. The/Every/No reading of the book is followed by a drinks reception.
       c. *The not reading of the book.
       d. *The having read of the book.
       e. *The being certain of the answer.
       f. John's loud singing of the song.
       g. *John's loudly singing of the song.
       h. John enjoyed a reading of the book.

It also seems to be the case, in general, that not only do Ing-of nominals allow determiners or quantifiers, they, in fact, require them – Hence

(6)   ?John enjoyed reading of the book.

where the book in question is the object and not subject of John's reading (i.e. John is not reading a book review). However, this form still remains somewhat productive as can be seen from a quote from a recent L.A. Times article , which quotes a judge as saying:

(7)   "There has been intentional withholding of documents and evidence in

---

[5]Internal syntax is the ability to be modified by adverbs or adjectives, to take determiners or quantifiers, control properties, etc. This contrasts with external syntax, which includes distribution in sentences, as well as agreement, scope and pied-piping.

> this case and . . . there's been intentional destruction of documents."[6]

The above sentence actually seems better than it would be with the definite article and this might have to do with it being a 'there' construction – later it will be shown that the quantifier 'no' can also occur with verbal gerunds in a 'there' construction. But in normal circumstances, this form does seem to be rather ungrammatical. Imagine the judge had expressed his personal reaction to the situation:

(8)   a. ?I am utterly shocked by intentional withholding of documents in this case.
        b. I am utterly shocked by the intentional withholding of documents in this case.

I bring this seemingly marginal example up because these structures (that I will refer to for convenience as [DET-] Ing-of) play an important role in upcoming sections, when it is shown that historically, things were very different.

Syntactic Analysis

So we are left with a rather strange situation as to the syntactic status of the various nominals. Ing-of nominals are perhaps the most clear, given that both externally and internally they behave exactly like NP's, and the construction apparently involves a straightforward deverbal noun, akin to derived nominals in phrases like *the destruction of the city.* But the verbal gerunds are certainly more puzzling. They have the external distribution of a noun phrase, but internally they behave like verb phrases (or clauses). However, it should be noted that even within the class of verbal gerunds, there are degrees of difference. Indeed, aside from the distribution facts, Acc-ing gerunds behave almost exactly like clauses (i.e. that-clauses, etc.), while Poss-ing gerunds are more of a mixed bag. This leads (Reuland, 1983) to classify the Acc-ing as a clause, while for Abney, its external distribution is enough for it to count as an NP (DP), with the clausal properties taken care of in the structure (more about this later). Anyway, it should be useful to point out a few differences between the Acc-ing and Poss-ing gerunds.[7] These examples all come from Abney.

   Agreement: With Poss-ing, two conjoined gerunds trigger plural agreement on the verb (like NP's), but this is not so for conjoined clauses or Acc-ing.

---

[6]L.A. Times 11 July 2001,"Judge Rules Against MTA Contractor".

[7]Their seems to be no settled position on the status of PRO-ing, i.e. is it a derivative of Poss-ing, Acc-ing or both, depending on context? Abney also suggests it may have a different structure altogether (but leaves it at that). I will take the standard view and consider them to be a special case of either.

(9)   a. That John came and that Mary left bothers/*bother me.
        b. John coming (so often) and Mary leaving (so often) bothers/*bother me.
        c. John's coming and Mary's leaving *bothers/bother me.

Long Distance Binding: Long-distance binding of subjects is possible in noun phrases and Poss-ing gerunds, but not with clauses and Acc-ing gerunds.[8]

(10)   a. they thought that each other's giving up the ship was forgivable.
        b. ?*they thought that each other giving up the ship was forgivable.
        c. they thought that each other's desertion was forgivable.
        d. ?*they thought that for each other to desert would be forgivable.

Pied-Piping: Again, where the gerunds contain wh subjects, whether or not they can front under pied-piping depends whether the gerund is Acc-ing or Poss-ing. And, of course, Poss-ing patterns with noun phrases, and Acc-ing with clauses.

(11)   a. the man [whose flirting with your wife] you took such exception to
        b. *the man [who flirting with you wife] you took such exception to
        c. the man [whose opinions] you took such exception to
        d. the man [(for) who to leave early] you would have preferred

Abney gives a number of other syntactic differences (some of which are rather tenuous), but this should be enough to see that they need to be given slightly different syntactic structures to account for their slightly different properties. However, (Hamm & Lambalgen, 2003), in giving a semantic account of nominalization claim that these syntactic differences have little impact on the semantics, and that there seems little semantic difference between the two types.[9] And indeed, it is quite plausible since *John's singing the song bothers me* and *John singing the song bothers me* seem to be synonymous. But it should be noted that there is an area in which there does seem to be a semantic difference – scope.

---

[8]I insert this bit for a bit of completeness in showing the differences between Poss-ing and Acc-ing. For myself, I am unsure of the grammaticality of either sentence in (10)-(a) and (b).

[9]They do bring up Reuland's example regarding distributional properties of gerunds with sentential adverbs (like probably) vs. VP-adverbs. Reuland's claim is that Acc-ing can take sentential adverbs, whereas Poss-ing cannot (this would be a difference in a semantic analysis).

(i)   a. John probably being a spy, Bill though it wise to avoid him.
        b. *John's probably being a spy made Bill think it wise to avoid him.
        However, Abney points that Acc-ing gerunds do not take sentence adverbials when in argument position:
        c. *I was grateful for John fortunately knowing the answer.

And as this data seems rather unclear, (a) and (b) are certainly not a big justification to treat the Poss-ing and Acc-ing gerunds different semantically.
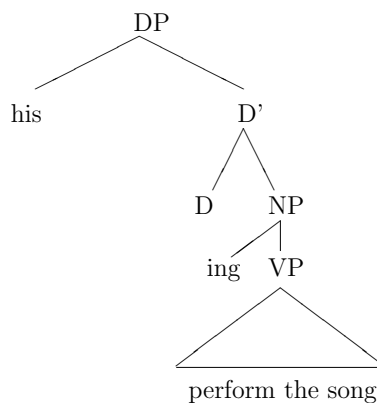
Figure 6.1: Poss–ing

As with the above examples, Poss-ing gerunds (like noun phrases) can take wide scope, while Acc-ing gerunds strongly prefer narrow scope:

(12)   a. John disapproves of everyone's taking a day off (wide scope OK)
         b. John disapproves of everyone taking a day off (*wide scope)
         c. John disapproves of everyone's happiness (wide scope OK)
         d. John prefers everyone to take a day off (*wide scope)

In (12-a) there is the reading where for anyone John disapproves of them taking a day off individually. Semi-formally we can see the wide scope as $\forall x[\text{Disapprove(John}, \text{take a day off}(x))]$. However, with (12-b), only narrow scope is possible, that is John disapproves of everyone taking a day off at the same time (a much more reasonable desire) – Disapprove(John, that $\forall x$(take a day off(x))).

It is now time to briefly examine some theoretical accounts (structures) of the various nominals. Abney accounts for the structure of all four gerunds in a modified version of X-Bar theory similar to (Jackendoff, 1977)'s

Deverbalizing Rule Structure: $Xi \rightarrow$ af – Vi

where Poss-ing gerunds are instantiated with X=N (category), i = 2 (bar level) and af = -ing (ing affix). This yields the following structure (for *his performing the song*) as can be seen in Figure 1:

For the Poss-ing gerund, this manages to account for the occurrence of a
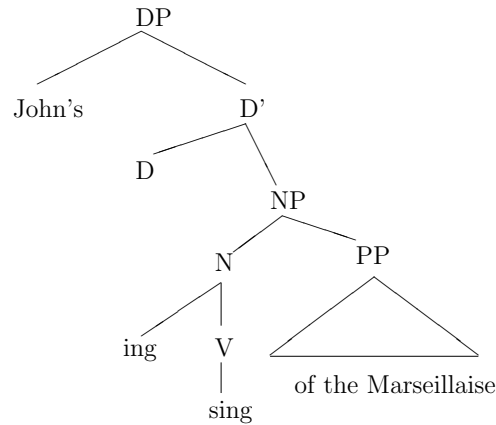
Figure 6.2: Ing–of

genitive, the presence of a VP (which is needed for the internal verbal properties – V" is a VP), the absence of a nominative subject, modals (**John's canning fix the boat*), and sentence adverbials like *probably* and *perhaps*, which are daughters of V"'(S) – which is not present in the structure.

Abney's version fully generalises to all the gerunds and he adopts a DP structure for the noun phrase. Again, it is the spot where the *-ing* is affixed that determines the internal syntactic differences noted above (as all structures are DP's, the external distribution is taken care of). In all cases, the *-ing* affix converts a verbal projection into a nominal one, but has no syntax of its own. For example, in the Acc-ing example, the *-ing* affixes to the IP and turns it into a DP, thus substituting its [+N] into the IP, converting it into a DP. Now as there is no D' or D, there are no nominal features to the structure other that its external distribution. As the *-ing* attaches lower (first Poss-ing, then Ing-of), there are increasingly more nominal features. For example, the Poss-ing structure has a D (determiner) position. According to Abney, the determiner position is the site of person, number and gender features, and so this accounts for the differences between Poss-ing and Acc-ing in terms of agreement as seen in (10). Note that in the Ing-of structure the *-ing* attaches to V0, yielding an ordinary noun that takes determiner, quantifiers, etc. The *Ing-of* version can be seen in Figure 2.

(Pullum, 1991) gives a rather simple structure in GPSG for the Poss-ing gerund (simple in the sense that no recourse to 'affix hopping' or morphological transformation is called upon). He gives the following rule:
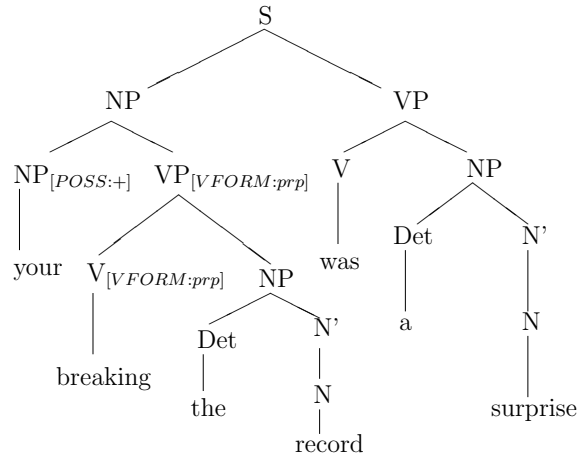
Figure 6.3: Pullum's NGP

(14) N[BAR:2] → (N[BAR:2], POSS:+]), H[VFORM:prp]

This is GPSG notation for indicating an NP structure that has a verbal (in this case participial) head, which is to account for the POSS-ing's external distribution as a noun phrase, and internal syntax behaving verbally. Note that bar level is treated as a feature of a category, and that H[VFORM:prp] (prp=present participle) means that the head of the noun phrase is [V:+, N:-], a violation of the head-feature convention (HFC) in which the features of the head daughter are inherited from the features of its mother (in this case NP and so [V:-, N:+). However, as this is a default, and not a hard constraint, a VP as head of an NP is allowed.[10] Also note that in Pullum's analysis both the feature [VFORM: prp] and [POSS:+] (indicating that the subject of the gerund (NP) is possessive) are not the result of some sort of transformation or affix, but assumed to come out of the morphology of the grammar itself. Pullum's structure is illustrated in Figure 3.

---

[10]Indeed it seems that one of Pullum's primary motivations in writing the paper was to show the usefulness of have the convention as a default (while still remaining meaningful). It should be noted that while there is a feature difference in [V:-, N:+] to [V:+, N-] from mother to head daughter, the bar level, amongst other features, is passed down to the head [Bar:2], in compliance with the HFC.

6.1.2 Semantics

(Vendler, 1968) distinguishes between two types of nominalization – Perfect and Imperfect nominals. Perfect nominals are a rather homogeneous semantic class that contain just the Ing-of nominals and semantically appropriate derived nominals such as "the destruction of the city."[11] On the other hand, Imperfect nominals contain the rest of our four gerunds (Poss-ing, PRO-ing, Acc-ing) as well as a host of other syntactic constructions like that-clauses, infinitival complements, and a few others.

The different distribution patterns of Perfect and Imperfect nominals can be seen by their occurrence or lack thereof in what Vendler calls Narrow and Wide containers. Narrow containers are verbal contexts, which accept only Perfect nominals, whilst Wide containers can accept either. Wide containers are contexts such as is unlikely, disturbed us, predicted; Narrow containers include occurred, was slow, was skilful. (13) and (14) give examples of nominals in Wide and Narrow containers, respectively:

(13)   a. John singing Karaoke surprised us.
        b. The performance of the play was surprising.
        c. The collapse of the Mid-East peace talks is depressing.
        d. Reading War and Peace made John feel like an intellectual.
        e. The singing of an encore is unlikely.

(14)   a. The destruction of the bridge occurred last week.
        b. The soprano's singing of the aria was too slow.
        c *The soprano's singing the aria was too slow.
        d. John's fixing of the sink was skilful.
        e. *John's fixing the sink was skilful.

While Perfect nominals are felicitous in either type of container, Imperfect nominals are only grammatical in a narrow container context. From this, it is quite easy to see a broad category distinction between the two types of nominals – events can be taken as the meaning of Perfect nominals, whilst facts or propositions can be taken as the meaning of Imperfect nominals.[12]

Indeed, for the most part, the gerundive Imperfect can be seen as more or less synonymous with *that*-clauses[13] or infinitival complements. Thus "John singing Karaoke surprised us" basically means "That John sang Karaoke surprised us.", and "Singing the song is fun" just means "To sing the song is fun". And we can

---

[11]Semantically appropriate in the sense the derived noun is 'close enough' in meaning to the corresponding verb. For instance, a performance is an event where someone performs something; however, an encumbrance is no an event where someone encumbers something – it is a thing which encumbers.

[12]Abney calls these 'act' and 'fact' readings respectively.

[13]Or often, clause that begin, 'the fact that'.

clearly see the event readings in (14), where, for example, in (14-a) there is a reading as to when the event of the destruction of the bridge occurred. And of course, *That the bridge was destroyed occurred last week* is pure nonsense. In (14-d) it is the way in which John fixed the sink that was skilful, and so a reference to the event of John's fixing of the sink. (14-e) could be felicitous if 'was skilful' refers to a strategy John has, for example, to avoid his wife becoming angry and fixing it before she found out. But in this case, it is the fact that he fixed the sink that was skilful.

But then what of the situation where a Perfect nominal (something that denotes an event) occurs in a Wide container? According to Vendler, Perfect nominals in wide containers can be interpreted as imperfect. Thus:

(15)   The collapse of the Germans is unlikely.

   is taken to mean

(16)   That the Germans will collapse is unlikely.

In some cases the Perfect nominal is coerced into having an imperfect reading in the context of a Wide container. But it should be noted that some containers are wider than others. In the case of Wide containers like *is unlikely*, and *predicted* (as in "Nostradamus predicted the collapse of the Germans") there is only a factive reading allowed. But take:

(17)   The Soprano's performance of the aria was surprising.

Here, there are two possible readings. One is the imperfect reading which is more or less equivalent to *That the Soprano performed the aria was surprising*, perhaps in a case where she had previously swore never to sing it again. The second is where it is something about the performance itself that is surprising – perhaps she sang it while simultaneously riding a unicycle and juggling. Indeed, it seems that context will help determine whether we get an 'act' or 'fact' reading (or possibly both) in these extra-wide containers. For example, look at:

(18)   a. The anti-globalisation demonstrations in Seattle, 1999 were surprising.
         b. The anti-globalisation demonstrations in Genoa, 2001 were surprising.

With (18-a) it is quite easy to get a 'fact' reading, because that there were large-scale demonstrations was quite surprising to most people. But by Genoa in 2001, one would have to be doing their best ostrich imitation to be surprised by the fact that there were large protests. Instead, it was the event itself that was surprising, (i.e. the scale of violence, police brutality, etc.) at least to some people.

Another interesting example of how context can affect readings comes from Abney in his discussion of N'-deletion (NP deletion for Abney) in gerunds. He

notes:

(19)   a. *John's fixing the sink was surprising, and Bill's [e] was more so.
        b. John's fixing of the sink was skilful, and Bill's [e] was more so.

(20)   a. John's fixing of the sink was skilful, and Bill's [e] was more so.
        b. *John's fixing of the sink was surprising, and Bill's [e] was more so.

In (19) both sentences involve the ellipsis of an NP, but only the (b) sentence is grammatical. Abney contends that NP-deletion is possible only under 'act' readings. And since the Poss-ing gerund can only have a 'fact' reading, (19-a) is predicted to be ungrammatical. But then what about (20-b)? While it is an Ing-of nominal, it is ambiguous as to whether it has an 'act' or 'fact' reading. Since (19-b) is a Narrow container context, the nominal must have an 'act' reading, and NP deletion is possible with 'act' readings. In (20-b), the default reading does seem to be for a 'fact' reading, as one is more likely to assume that it was the fact that John did fix the sink was surprising. According to Abney, NP deletion is not possible with 'fact' readings, and so the lack of ellipsis is in fact unsurprising. But, an 'act' reading, where it would be the way the sink was fixed was surprising is possible with a little imagination. However, I invite the reader to imagine the admittedly bizarre situation, where John fixes the sink merely by putting all the parts on the floor and uttering 'Abracadabra'. Bill does this as well, but also conjures the plumbing parts out of midair. Now with this in mind, (20-b) seems perfectly fine, there is now an 'act' or 'manner' reading.

The lesson here is that it is a combination of context and lexical meaning/world knowledge that determines whether a Perfect nominal in an extra-wide container like is surprising gets either an 'act' or 'fact' reading. While the 'act' reading in (18) comes out quite naturally (because the manner of demonstrations or performances and the like are often surprising), John's fixing of the sink is, by default, probably unremarkable in the way it is done. Indeed, a rather ridiculous context needs to be constructed in order to make (20-b) sound reasonable.

Before concluding, one last difference between Perfect and Imperfect nominals should be noted – intensionality vs. extensionality. Take the following examples from (Parsons, 1990):

(21)   The beheading of the tallest spy occurred at noon.

(22)   Mary predicted the beheading of the tallest spy.

Now assume it just so happens that the tallest spy is actually the king. Then (21) implies:

(23)   The beheading of the king occurred at noon,

but, (22) does not imply

(24)   Mary predicted the beheading of the king.[14]

At first glance it appears that Narrow containers enforce extensionality, but wide containers do not; however, it seems more likely the case that it is the 'act' reading that forces extensionality and not the container per se. Narrow containers, only allowing 'act' readings, would then be a special case. Now in the case of the sentences with 'predicted', we have a container that, as noted earlier, only allows for 'fact' readings. But with a container like 'shocked the world', which allows both readings, extensionality is enforced for 'act'/event readings. Consider:

(25)   The bombing of Hiroshima shocked the world.

Presumably, the default reading here is the event reading, it being the first use of an atomic bomb that was most shocking (i.e. not just that Hiroshima was bombed). Now, assume that Hiroshima was the fourth largest city in Japan at the time. Then certainly (25) implies

(26)   The bombing of the fourth largest city in Japan shocked the world.

As far as the 'event' reading goes, an event by any other name will still shock as much. Of course the 'fact' reading of (26) does not go through on either reading of (25). So perhaps it is whether the interpretation of the Perfect nominal is as an event or propositional entity that determines the extensionality. If it is an event, then extensionality is enforced, if it is a fact, then the context is intensional.

Generic Nominals

There is one class of verbal gerunds that are commonly held to have an activity reading, and are a variant of the PRO-Ing gerund, where the subject of the gerund is no specific subject for the determinerless gerund. (Lees, 1960) called these action-referent gerundives,examples of which are

(27)   a. Eating vegetables is healthy.
          b. Dressing oneself is fun.

These generic nominals differ from the types discussed above in that they are not paraphrased by a *that*-clause, but rather an infinitival.[15]

(28)   a. To eat vegetables is healthy.

---

[14]It is interesting to imagine the situation where Mary knows that the king is the tallest spy. With this added assumption, does the implication of (24) go through? Suppose Mary utters something along the lines of *I foresee that the tallest spy will be beheaded tomorrow*, but purposely leaves her secret information out. For herself, at least, (24) may be the case. Substitute 'believed' for 'predicted' and it certainly seems to go through.

[15]The use of an Ing-of gerund is also allowed, with no interpretational difference, i.e. *The eating of vegetables is healthy.*

        b. To dress oneself is fun.

When these determinerless gerunds have an understood, specific subject, the interpretation supposedly reverts to a 'factive' reading, with the paraphrase being a possibly a *that-clause*, and sometimes equivalent to a (tensed or untensed) infinitival.

(29)   a. I regret dressing myself.
        b. I regret having dressed myself.
        c. I regret that I (have) dressed myself.
        d. *I regret to dress myself.
        e. *I regret to have dressed myself.

(30)   a. Climbing the mountain was quite satisfying.
        b. ??That I climbed the mountain was quite satisfying.
        c. To climb the mountain was quite satisfying.
        d. To have climbed the mountain was quite satisfying.

The context in (29) is undoubtedly factive and its corresponding paraphrases do not seem to allow an infinitival. However, while the context in (30) is both a specific event and the subject of the gerund is the speaker, there is certainly an interpretation that it was the action of climbing and not merely the result that is satisfying. It is also much more amenable to an 'actional', infinitival paraphrase.

Conclusion

In this introductory section, we have seen some basic syntactic and semantic properties of gerunds. Syntactically, all four types of gerunds can be taken to be NP's because of their distribution. However, only the *Ing*-of gerunds can be fully nominal, as the other three gerunds possess internal syntactic properties of verb phrases in their ability to be modified by adverbs, their lack of ability to take quantifiers and determiners, and the control properties for PRO-ing.

    The syntactic differences between the verbal and nominal gerund seem to be exhibited semantically since the verbal and nominal gerunds belong to the classes of Imperfect and Perfect Nominals respectively. The 'fact' and 'act' readings that characterise the semantic differences between the Imperfect and Perfect Nominals can be taken to denote propositions/states of affairs and events.[16] Many syntactic and semantic subtleties have been excluded here, but I think this provides a firm foundation on which to examine the diachronic status of nominalization. We shall see a rather strange evolution of the gerund from its origins as a simple

---

[16]The relation between 'event' and 'act' is fairly tricky. As seen above, events can have 'act' readings in wide container contexts and certain narrow containers like 'is slow'. But how does this relate to the habitual 'act' readings such as in *John's singing of arias is invariably out of tune*? No specific event is implied here, but perhaps the John's quantifies over singing-aria events of John's and states something of them in a default way.

deverbalized noun with a resultative, concrete interpretation, such as *clothing* or *building.*

6.2   Origins (and present-day consequences)

In the previous section, a categorical difference between the nominal and verbal gerund was shown. The nominal types allow determiners and mark their direct object periphrastically with *of.* The verbal gerunds[17] allow modification of aspect and voice, and take a direct object without any recourse to periphrasis. Finally, it was observed that the semantic objects the nominal and verbal gerund denote also differ.[18]

   This section examines the history of the nominal, as it is important to recognise that the differences between the nominal and verbal gerund were not always so clear. Indeed, *-ing* forms used regularly as late as the end of the 19th Century are startlingly ungrammatical today.[19] Following our historical excursion, we shall see that even in PDE,[20] the discrete categorisation given in the last section proves wanting. Moreover, the standard mapping between syntax and semantics, as well as the standard typology of nominal types needs to be revised slightly.

6.2.1   Old English and early Middle English

Until 1300[21] verbal gerunds were yet to exist, but there was a productive derivational process that turned a verb into an abstract action noun[22], via the addition of the suffixes of *-ing* or *-ung.*[23] Here are some common examples of Old English verbs and their corresponding verbal noun.

(31)   a. *gieddean* to speak formally, with alliteration

---

[17]Recall that (Abney, 1987)'s typology includes three verbal gerunds that do differ slightly in their syntactic behaviour with respect to each other.

[18]Excluding, of course, the coercion of perfect nominals to a factive reading in the proper context. In later sections, we shall see that this is certainly not the only type of coercion possible for gerunds. Indeed, while a less regular phenomenon, a type of coercion is possible, where verbal gerunds have a perfect nominal reading, and often some nominal syntax to go along with it.

[19]Though (Pullum, 1991) contends quite reasonably that latter examples of bewildering gerunds may be an instance of deliberate archaisms by late 19th and early 20th century authors.

[20]This is often dialect dependent.

[21]The standard division of periods of English is usually divided into

Old English	500-1100
Middle English	1100-1500
Early Modern English	1500-1800
Modern English	1800-Present

[22]Or a concrete noun such as still exist in PDE, i.e., *I love the clothing she wears.*

[23]Or their inflected forms *inge* or *unge.*

     b. *gieddung* saying
     c. *lufianto* love
     d. *lufung* act of loving
     e. *geladian* invite, summon
     f. *geladung* congregation[24]

These early *-ing* forms had only nominal properties, needing a periphrastic direct object, and not exhibiting tense or voice distinctions. (Tajima, 1985) has a typology of gerunds of Old and Middle English[25]

(32)   **I** objective genitive + gerund: (the *kinges couroning*; Bevis, c1300)
    II object + gerund: (be other *pennaunce doynge*; Rolle)
    III gerund + of-adjunct: (the beginning of wysdome is dredyng of our Lord; Midland Prose Psalter, c1300)
    IV determiner + gerund + of-adjunct: (Ye han wel herd of Theseus devyse in *the betraysynge of fayre Adrayne*; Chaucer, c. 1380)
    V gerund + object: (In *baptising bath yong and ald Men soght til him* (Cursor Mundi, 1348)
    VI determiner + gerund + object: (*The wythholding you from it* can doo yow no good; Caxton, c1481)

The focus shall be on types **III** to **VI**.[26] Notice that Types **III** and **IV** differ from types **V** and **VI** only in the presence or absence of an *of*-adjunct. Type **VI** is considered rather ungrammatical today[27] but for several hundred years was a rather common before being dying out for a combination of prescriptive and, according to (Fanego, 2004), language internal reasons. Type **V** is an early example of an ancestor of the PRO-*ing* gerund, which was the first type of verbal gerund to emerge. The earliest extant example of this is

(33)   Sain John was...bisi in ordaing of priestes, and clerkes, And in *casting kirc werkes* (English Homilies, ca. 1300)

---

[24]This last one from (Houston, 1989, p.174).

[25]His types III and IV are the most relevant for the development of V and VI. Note at the time of his study, ACC-ing gerunds were yet to exist.

[26]See (Tajima, 1985) for greater detail on the first two types, which are interesting in their own right.

[27]It combines a determiner with a non-periphrastic direct object, giving a nominal/verbal hybrid.

which is a coordination of two type **III** gerunds and the then novel Type **V**.[28] It took several centuries for the various modern verbal types to completely develop, and along the way some (from our perspective) rather odd forms can be found. This brings up the questions as to how it happened and why the process was so gradual. In fact, it may not yet be complete.

The table below[29] is taken from (Tajima, 1985) and is a compilation of gerund type and percentage over 50 year periods. He has a few counts for some very early appearances of the verbal gerund's ancestors. This is not uncontroversial. For example, many of his examples include *ing* forms with certain adverbial modifications that could also modify nouns.[30] Nevertheless, the table does show a gradual expansion of the 'verbal gerund' both in frequency and variety of types. Apparently, the Acc-ing gerund really doesn't appear much until the late 16th century and not with regularity until the late 18th century, and is well beyond the time period covered by Tajima.

	I	**II**	**III**	**IV**	**V**	**VI**	Total
1100-1200	12	9	23	8	0	0	52
1200-1250	22	9	3	9	0	0	43
1250-1300	17	22	19	5	1	0	64
1300-1350	82	51	160	103	23	2	421
1350-1400	35	60	595	250	59	4	1003
1400-1450	82	227	717	414	253	7	1700
1450-1500	74	102	562	594	328	17	1677
Total	324	480	2079	1383	664	30	4960

6.2.2 The OE participle and its role

The standard view of the development of the verbal gerund is that a confusion developed between the Middle English present participle and the verbal noun. As the endings of OE verbal nouns coalesced from *ung* and *yng* to *ing*, participial endings underwent a similar phonetic change. Depending on the dialect,[31] the

---

[28]Tajima has some examples of what appear to be earlier verbal gerunds – mainly to do with different types of adverbial modification such as place or time, i.e. *of thy comyng at domesday* (1280), *at his coming there.* (Fanego, 2004) notes that prepositional phrases and phrasal adverbs could also occur with nouns at the time (and still can), thus making their status unclear. However, she does acknowledge that the increasing use of certain types of adverbial usage expanded in Middle English may have contributed to a categorial indeterminacy that is central to her thesis.

[29]Given the relative poverty of data from the 1100-1300 time period, it is probably unwise to draw any conclusions regarding trends in frequency for that period.

[30]See (Fanego, 2004) for more details.

[31]i.e. Southern, Midland, Northern, etc.

present participle had an ending of *inde*, or *ende* or *ande*.

fremmann to perform
fremmende performing

nerian to save
neriende saving

(34)   He waes ehtende   cristenra mona
         He was   persecuting Christian men

(35)   Tha   waeron simbe binan Romebyrg wuniende.
         Those[senators] were   always within Rome   dwelling.
         Those [senators] were always dwelling within Rome.[32]

(36)   Swete lord... Ich am cominde to thine feste. (c1280)
         Sweet lord... I   am coming   to thine feast.

(Poutsma, 1923) then describes a process of the levelling of the participle in South and Midlands dialects (which had the *inde* ending) first to *inne*. It is then claimed that this became conflated with the verbal noun ending in inge . And indeed, in the South and Midlands dialects the participial and verbal noun endings were identical by 1450 (Poutsma estimates the beginning of this process to be around 1200). Here are two examples of early examples of participles ending in *ynge/inge*.

(37)   And how louynge he is to ech   lif   on londe and on water
         And how loving   he is to each living-thing on land and on water
         (c1400, Piers Plowman)

(38)   Jhon was in desert baptisynge and   prechinge the   baptym
         John was in the   desert   baptising, and   preaching the
         of   penaunce. (1380,   Wyclif)
         baptism of   penance.

There is now a verbal noun and participle that sound indistinguishable (in the South, at least). The conflation theory then goes on to postulate that it was this phonetic formal identity between the participle and the gerund that caused the gerund to acquire verbal properties. The participle is verbal, and takes bare direct objects, passivizes, governs a predicative complement, etc:

(39)   I recommand me hartly onto yow, thankyng yow of aull good brother-

---

[32]This and (35) taken from (Denison, 1983) in the context of analysing the OE progressive. I am using them here only to show the PP with the *ende* form.

> hood. (Cely Papers, late 1400's)

(40)   Unto my brother George Cely merhcande of the estapell being at Calles.
         (Cely Papers)

And apparently, speakers of English got confused by phonetic identity, and the
gerund started to take on the verbal properties of the present participle. Of
course, the timing of the formal identity (ca. 1450) does fit in nicely with estimates
of the beginning of the verbal gerund (ca 1350-1450),[33] but some questions do
remain.

First, while there was the formal identity of the present participle in the
South and Midlands, even Poutsma admits that in the North and in Scotland,
the distinction still remained. In fact, it is beyond dispute that Scottish English
had a verbal gerund by 1560, but still kept the old participle ending. These
examples come from St. Andrews, Scotland court records involving accusations
of blasphemy:[34]

(41)   Margaret murdow delatate for blashphemous sayings against the Sacra-
         ment of the body and Blude of Christ *sayand thir wordes* in the oppin
         fische mercat...

(42)   William petillok *dwelland be este Thomas martynes* Said the divell ane
         kirk will I gang to and the devil burn up the kirk...

(Labov, 1989) notes that the formal identity may never have taken place in
some dialects. In examining the socio-linguistic variable in PDE dialects between
workin' and *working* he notes that where the former is spoken (weakened ending),
it is used most in progressives and participles, rarely in gerunds, and least of all
in deverbal nouns like ceiling. He takes this as evidence that the workin'/working
alternation is not the result of the deletion of the underlying /g/, but rather a
continuation of the distinction in some dialects between the participle and the
gerund. The *workin'* form comes from a reduction in the *inde* participle.

So, if the formal identity of the participle and gerund was responsible for
the split of the gerund into a verbal and nominal one, then how did this hap-
pen in regions where the identity never actually occurred? Furthermore, if we
grant a formal phonetic identity, as Houston wonders, is it sufficient for "confu-
sion between two grammatical categories, unless there were already some shared
grammatical properties"?

6.2.3   Houston

(Houston, 1989), came up with a novel proposal for the origin of the verbal gerund.

---

[33]The first example of an unequivocal verbal gerund is from 1300, but no great category
change happened immediately. It is over the next 100-200 that they become more frequent.

[34]Examples taken from (Cusack, 1988).

She begins by questioning the standard theories that rely on formal phonetic identity, noting first that the change in the participle ending form *nde* to *ing* was only completed by 1450 in the south of England and the Midlands.[35] Moreover, she questions the possibility of a 'confusion' between two different grammatical categories, unless there were also some shared grammatical/syntactic properties. She cites data from (Irwin, 1967) that shows that verbal nouns occurred primarily as subjects, objects, objects of prepositions, and genitive complements to a lesser extent. Participles occurred mainly as nominal modifiers and marginally as parts of phrasal verbs. So as far as external syntactic distribution, there were only a few similarities between verbal nouns and present participles (and this remains so in PDE).

But Houston does not abandon the idea of the participles involvement in the evolution of the verbal noun altogether. Instead, she looks at the possibility of a shared discourse function between the two, and theorises that it is this shared communicative function that caused the gerund to split into a nominal and verbal gerund.[36] Her data does contrast rather sharply with Tajima's, as she only finds gerunds taking bare direct objects in any frequency by 1550, about 150 years later than Tajima, and 100 years after we see gerunds with direct objects appearing in letters and original English prose. I have no explanation for this rather large discrepancy,[37] but as we shall see, a key part of her data (the percentage of these direct object gerunds that appear as objects of preposition) is consistent with Tajima.

She first divides up the occurrence of verbal nouns by syntactic position and contrasts whether they are found as subjects/direct objects or objects of prepositions. Her data ranging from the 10th to 17th centuries shows that from the 14th century on, at least 60% of all verbal nouns are found in oblique position. This does agree with Tajima, as his examples in all of his types have the gerund in oblique position in a fairly large majority of the instances. Houston then looks at what happened when the verbal noun began to take bare direct objects. It turns out that the first appearance of bare direct objects is with verbal nouns in the oblique position.

---

[35] As already discussed in the last section.

[36] Houston labels these 'oblique' and I will continue this trend as it is briefer than 'object of preposition'. I am using the term 'verbal noun' to mean any gerund. 'Direct object gerunds' are gerunds that take a bare direct object (Type V) and are of course looked at as being the first hints of the development of the verbal gerund.

[37] Indeed Tajima's timing of the beginning of the change is probably more accurate. But, while there are unequivocal examples of verbal nouns with direct objects as far back as 1300, there are a very rare occurrence (as far as what is extant) for over a century.

c1550   4% of sub/obj position w/ direct object
c1600   54% of sub/obj position w/ direct object
c1650   38% of sub/obj position w/ direct object

c1550   31% of oblique position w/ direct object
c1600   64% of oblique position w/ direct object
c1650   60% of oblique position w/ direct object

The first thing to note is the 50-year gap between when direct objects first appear in oblique contexts and when they really begin to appear in subject or object position.[38] While Tajima does claim there to be direct object verbal nouns[39] in rather high frequency appearing from 1400, it seems that every one of them he gives as examples also occur in oblique position. For example, he lists at least 25 examples for the 1450-1500 time period. None of them are in either subject or object position. So despite the disparities in the data, it does seem that indeed the first occurrences of verbal nouns taking bare direct objects did occur with the verbal noun in the oblique position.

The dominance of verbal nouns in oblique position at this time is important; Houston claims that there are shared discourse functions between the oblique position verbal noun and the appositive participle. Then it might be the case that the shared discourse functions caused the gerund to acquire verbal properties. If this did happen, one would expect to find bare direct objects with verbal nouns in oblique position before they appeared in subject or object position. And this is indeed the case. So we shall now have a brief look at the appositive participles and the shared discourse function between them and oblique verbal nouns.

According to Houston,[40] appositive participles appear either before or after the matrix clause, or in clause-final position:

(43)   a. *Going to preach*, H. Morley of my parish deliv'd mee a note of receipt of my procuration (Diary of Ralph Josselin) b. I recommand me hartly onto yow, *thankyng yow of aull good brotherhood*.[41] (Cely Papers)

Like verbal gerunds the subject of the appositive participle exhibit the same sort of control properties. The subject of *thankyng yow* in (43) (b) is the subject of the

---

[38]Unfortunately her data is not quite as copious as Tajima's; for example the 4% for the time period 1550-1600 represents but one instance out of a total of just 24 verbal nouns found in sub/obj position at all. This may be the reason for the peak at 1600. For the 1650 time period, there are but 13 examples of verbal nouns in sub/obj context and 76 verbal nouns in total.

[39]The unwieldy term 'direct object verbal noun' is used to classify examples such as (56). Certainly at this earliest date they cannot yet be called verbal gerunds, but rather the beginning of the transition stage that saw verbal nouns begin to acquire internal verbal syntax.

[40]Appositives, in general, are clauses that re-identify the subject of the matrix clause such as *That man, the president of the corporation, is out to sack me.* For example, the participle in *Feeling tired, I went to bed* would be appositive.

[41]Note that all data in this section is taken from Houston.

matrix clause ('I'), and in (43-a), the subject of 'going to preach' is the speaker, presumably Ralph Josselin.[42] And as far as taking bare direct objects goes, she cites (Callaway, 1901) in noting that at least in Old English texts the appositives took direct objects in 56% of their instances in prose and 13% in poetry.[43]

She also draws on Callaway for her theory of similarity of discourse function. According to Callaway, appositive participles serve three different functions, attributive, adverbial and coordinate:

(44)   Unto my brother George Cely merchande of the estapell *beyng at Calles* (Cely Papers – attributive use)

(45)   that lewide men (laymen), *seyinge akynge & swellynge in a lyme that is wounded*, leie therto a potage in maner maad of eerbis & Wynes...(Science of Cirurgie – adverbial[44])
        That laymen, *seeing aching and swelling in a limb that is wounded*, lay thereto a potage in manner made of herbs and wines...

(46)   The Quene removed on Wensday toward Norfolk, *taking Dr. Cesars in her way.* (Letters of John Chamberlain – coordinate)

Houston focuses on the adverbial function of participles, as this is the function shared by verbal nouns in oblique position (there was no adverbial function with verbal nouns in either subject or object position). Following Calloway, she distinguishes between several types of adverbial functions including manner/means, temporal, causal, and goal. Here are a few examples of participles and verbal nouns in some of the functions:

(47)   Sir Samuel Baguel is lately slain there, *being stabd by Sir Laurence* (Letters of John Chamberlain – appositive, causal function

(48)   God zelde yow for zoure labore for me for *gaderyng of my mony* (Paston Letters, verbal noun – causal function)

(49)   (He) set upon him as he was coming out of his coach, *wounding him in three or four places* (John Chamberlain – appositive, manner function)

(50)   Wee are very vigerous in *asserting our Religion* (Essex Papers, late 1600's

---

[42]It is a rather odd sentence indeed. The main clue that equates *mee* with the subject of *going to preach* is *my parish*. Nonetheless, the participle is controlled by the matrix clause.

[43]Once again Latin rears its imperial head. In the OE period, according to Callaway, most writing that is in original English is poetry. Most of the prose consists in Latin translations. And again a Latin influence is cited for the reason that appositive participles taking direct objects. Callaway also notes that the strictly the Latin translation the higher the frequency of appositives taking direct objects.

[44]This is adverbial in the sense that the participle bears a temporal/causal relation to the matrix clause. That is, it is upon and because of seeing the wounded limb that the laymen decide to do their attempt at surgery. I assume, since this is a treatise on surgery and laymen are the subject, that it all goes drastically wrong.

– verbal noun, manner)

Finally, data is presented showing a fair amount of similarity in the percentage of both appositives and verbal nouns in an adverbial function. For all but the period 1650-1700 the percentage of oblique verbal nouns with an adverbial function is well over 70%. Aside from a strange dip in the period 1450-1500, the percentage of appositive participles with adverbial function is at least 50%. In all time periods, the percentage of subject or object verbal nouns with adverbial function is 0.

Summing up, there exists a similarity in discourse function between appositive participles and oblique verbal nouns, and evidence that when verbal nouns began taking direct objects, it was the oblique ones that took them first. Houston speculates that the common usage, some syntactic similarity (ability to appear clause-initial and clause-final), and functional similarity between the appositive participle and verbal noun could have contributed to increasing similarity between the participle and gerund, to the point where the gerund began to acquire verbal properties.[45] And while this certainly has a good deal of intuitive plausibility, there is no broader theoretical mechanism to show how this would actually work. In the next section (Fanego, 2004) attempts to give a mechanism that could account for the glacial pace changes that happened to the development of *ing*.

There is also the discrepancy in data to take into account. But assume Tajima's data is correct and that verbal nouns began taking direct objects in meaningful frequency by the period 1400-1450. This is not enough to show that a verbal gerund was necessarily in English by that time – at least not the verbal gerund as we know it. It is a gerund generally without a subject,[46] that only has an external distribution in oblique position. There are also only a few examples of the gerund with a predicative complement, and one or two odd cases of any voice distinction. There is certainly no verbal gerund as it exists today. So perhaps the end of the time period that Houston looks at (ca. 1500-1600) is actually the time when a real split does occur, and the previous 150 years or so only some kind of transitional stage. By 1600, the verbal noun begins to take direct objects in subject and object position. Moreover, we also now start seeing in much larger frequency what look to be Poss-ing and Acc-ing gerunds, and distinctions for tense and voice also become established.

Before the late 16th century, the gerund was neutral as to voice:

(51)   a. A shootynge Gloue is chieflye for to save a mannes fingers from hurtynge (Aschams, 1545, = a man's fingers from *getting* hurt)

---

[45]Houston does not hold that this was the sole reason, just that it was a contributing factor. In a footnote she also draw attention to the fact that issues of control may be involved. Given the odd status of the [Det-] Ing-of gerunds, there is most likely something to this.

[46]Tajima's examples of gerunds with subject seem mainly nominal, and there are no examples of an Acc-ing like gerund, i.e. with both subject and direct object.

b. The witnesses said they dared not to present the truth for drede of murdrying, and to myscheved in their own houses (Rolls of Parit, 1475, = being murdered)

c. It please you to remember (remind) my maister at your best leiser, wheder his old promise shall stande as touching my preferrying to the Boreshed (Paston Letters, 1458, = being preferred)

But around the latter half of the 16th century the distinctions start appearing:

(52)   a. Cloth'd with a pitchy cloud for being seen (Marlowe, 1587)
b. to be styled the underhangman of his kingdom and hated for being preferr'd so well (Shakespeare, 1591)

and distinction for tense also came in:

(53)   a. Want of consideration in not having demanded thus much (Sidney, 1580)
b. Mine ears ... Do burn themselves for having so offended (Shakespeare, 1592)

but nevertheless the voice-neutral construction seems to be variable well until the 19th century, and was even used in the same dialect as the passive form

(54)   a. Excuse his throwing into the water (Shakespeare, 1598)
b. devoting all His soul's strength to their winning back to peace. (Browning, 1833)

Perhaps then, the period in question, 1500-1600, is not the time that verbal nouns began to acquire direct objects, as that seems to have started well before, but instead is the time there began to be a regular, productive construction of a verbal gerund, with the majority of features that are found in the verbal gerunds in Present Day English.

6.2.4   Fanego

Indeed, recent work by (Fanego, 2004) pinpoints the beginnings of the verbal gerund slightly further back than Houston, but later than believed by Tajima. Many of Tajima's earliest examples involve types of adverbial modification that was common to both verbs and nouns in early Middle English and therefore his earliest examples cannot be seen as definitively verbal. For example, Tajima considers this example from 1280 an example of a verbal gerund.

(55)   Vnder the þe Monument ȝeo stod wiþoute wepyng *sore*. (Southern Passion)
She stood close by the sephulcre without weeping bitterly/without bitter

weeping.

Citing (Jack, 1988), Fanego notes that *sore* at this time in history was (due to phonological and morphological processes) both an adjective and adverb, and thus could have an adjectival (and thus nominal) reading in the above example. In PDE, words such as *fast* still have this ambiguity.

The safest way of determining the verbalisation of the gerund is in its governance of direct objects,[47] which Tajima also has very early examples of:[48]

(56)   a. Saint Jon was...bisi In ordaining of Priests and clerics, and in casting
        kirc werkes (1300, English Metrical Homilies).
        Saint John was ...busy ordaining priests, and in planning church works.
        b. yn feblyng *þe body* with moche fastyng (1303, Handlyng Synne)
        in weakening the body by too much abstinence.

These two examples are probably the earliest clear examples of a gerund with verbal properties. Note that in the (56-a), there is a combination of the nominal Type III (*ordaining of priests*) with the then novel Type V (*casting kirc werkes*). Moreover, there is support of Houston's chronology in that both of these earlier examples are gerunds as objects of preposition.

Fanego's work differs from Houston not so much in the account of the chronology and development of the verbal gerund, but rather in the mechanism. She points out that while Houston is perfectly accurate in attributing the beginning of the process to oblique gerunds, the connection with appositive participles, though plausible, gives no mechanism to account for the *process* that takes several centuries to unfold.[49] (Fanego, 2004) does exactly this in terms of reanalysis and actualisation, and examines a piecemeal process that begins with the early Type V's seen above and gradually develops into the typology of verbal gerunds in PDE.

Reanalysis and Actualisation

The basis of Fanego's use of the term 'reanalysis' comes from seminal work by (Timberlake, 1977) and (Langacker, 1977). The simplest definition she gives comes from (Timberlake, 1977) that defines reanalysis as "the formulation of a novel set of underlying relationships and rules" (p.141) and a fuller quote from (Langacker, 1977) as

---

[47]i.e. *sans* an *of* marker.

[48]The first of these has already been seen in the section on Houston.

[49]Moreover, Fanego even dismisses Houston's attempted semantic and discourse link to account for the verbalisation process beginning with the oblique gerunds. Instead, more syntactic factors are relied upon. Later on, we shall see that Houston may indeed have a point; however, Fanego is correct in asserting that a mechanism to account for the rather complicated history does not need to be given.

> change in the structure of an expression or class of expressions
> that does not involve any immediate or intrinsic modification of its
> surface manifestation. Reanalysis may lead to changes at the surface
> level, ...but these surface changes can be viewed as the natural and ex-
> pected result of functionally prior modification in rules and underlying
> representations.

She further refines this with reference to (Harris & Campbell, 1995) and
(Croft, 2000a) noting that the major prerequisite for reanalysis involves a given
form having the possibility of more than one analysis. Moreover, it is not the case
that every example of a particular type is a candidate for reanalysis, a subset of
tokens is enough to begin the reanalysis process. As Croft notes, a construction
need not be formally ambiguous in order for there to be a remapping between
form and function.

(Fanego, 2004) then postulates that the first candidate for reanalysis were
Tajima's Type III gerunds[50] as indeed the earliest examples of gerunds with
direct objects are also lacking a determiner. She notes that in general, *ing* nouns
were a favoured candidate for a reanalysis into something more verbal due to
a variety of factors. First of all, by Middle English, *ing* nouns have unlimited
productivity and consistently having a meaning, 'act/process of doing X' as well
as preserving the argument structure of the base verb.[51]

At the time, Type III nominal gerunds were far more common than in PDE,
and that they were lacking a determiner means that they were not less-marked
as overtly nominal than those gerunds that did have a determiner.[52] In her view
it was *all* of these factors that allowed speakers to reanalyse the Type III gerund
as something more verbal. Once it is analysed as verbal, the *ing* form taking a
direct object follows immediately.

Note that this sort of reanalysis can happen even when the old analysis is
perfectly unambiguous. There need only be a possibility of multiple analyses –
old and new analyses can happily coexist. As seen in (56)(a), the writer had no
problem using a typical nominal gerund alongside the innovative verbal gerund.

Finally, her explanation for the reason that the first nominal gerunds to be
reanalysed were Type III in a prepositional context does not have to do with the
discourse reasons. It is rather the more prosaic reasons that prepositional gerunds
were the most common at the time,[53] as well as influences from the French '*en*
plus *gérondif*',[54] coupled with the resistance of the use of infinitives as objects

---

[50]Fanego calls these Type I (*Ing-of* with no determiner).

[51]She also cites various factors discussed earlier such as influence from other languages, ad-
verb/adjective ambiguity, etc.

[52]Nominal gerunds were not alone amongst nouns in having a zero article – this was much
more common in the grammar of middle English than today.

[53]Something that has already been noted.

[54]An old French prepositional construction that entered Middle English.

of prepositions, thus removing one alternate form that may have been used.[55] However, what neither of the authors have pointed out is that the verbal gerunds in (56)-(a) and (b) have a meaning that is in no way 'factive', but is exactly the same as its Type III counterpart – an act or process.[56]

Actualization and grammatical gradation

At this point, the story is that the Type III nominal gerund construction had enough syntactic properties that allowed it to be interpreted as verbal as well as nominal. But, we have seen in the history that the early examples from the 14th century did not immediately beget a PDE type verbal gerund. Instead, there was a gradual working out of various constructions, most of which are ungrammatical today.

So, instead of viewing grammatical categories as discrete, Fanego calls upon work of (Haspelmath, 1998) as well as (Heine, Ulrike, & Hunnemeyer, 1991) that views language change as a gradual process. They note that in many language change processes[57] there are often stages where a particular entity cannot be described as, say either a verb or noun, but a hybrid of both, as it retains properties of both categories. Over time the process works itself out, and the given entity or construction gradually becomes an archetypical member of the target category.[58] Viewing word classes in a more dynamic way makes the strange, hybrid anomalies that occur over the gerunds history seem not so odd at all, but something perfectly expected.[59]

In Fanego's account the early examples of gerunds taking a direct object are only the beginning of a process that then extended to other types of nominal gerunds in further extensions, called the actualization process. Extension is "the analogical transference of a pattern to further related cases." This idea is rather similar to many other author's conceptions of language change, including (Hopper & Traugott, 1993)'s *rule generalisation* and (Croft, 2000a)'s *intraference*. For Croft, *intraference* is the "extension of a form to a function not previously associated with that form" (p.154) on the proviso that there is enough meaning overlap between the form's original meaning and the new function it becomes associated

---

[55]As seen in the first section, some uses of gerunds, especially those without a subject can be paraphrased as infinitivals.

[56]Appositive participles also often have an action reading that serves as a background for the main clause.

[57]This is often typical in grammaticalisation processes where a verb becomes a preposition over a long period of time.

[58](Haspelmath, 1998) goes as far to say that even in formal syntactic representation it is probably correct to allow for word class gradience. At any point in time a word or construction could be categorised as, e.g., $V_{.2}/P_{.8}$ for a word that is basically a preposition, but retains some properties of a verb.

[59]Though, verbal gerunds are in general much more worked out and standardised in this point in time, a following section will show that the transition from noun to verb still has yet to finish completion.

with. In PDE, it is often the case that using a verbal or nominal gerund (or even derived nominal) is merely a matter of style,[60], and we have also seen that in the earliest examples of verbal gerunds the meaning is just the actional reading that is shared both by Type III gerunds and appositive participles.[61]

Further extensions

The next two stages in the development of the gerund involve the development of verbal features in Tajima's Type IV gerunds (*the/my writing of the book*) and the extension of Type V gerunds (*in casting kirc wirkes*) to non-propositional contexts.

The Type IV gerunds, especially those with a determiner are certainly more nominal and are, as Fanego claims, a less natural candidate for acquiring verbal properties. Fanego notes that even by the late 16th century, these type remained primarily nominal whereas those without a determiner or possessive were primarily verbal[62] by 1640.

Note that Tajima's Type IV include nominal gerunds with both a determiner and a possessive. The second type, which would develop into the modern POSS-ing gerund verbalised comparatively quickly, given their more clause-like nature. Astoundingly, Fanego does a sample of 400,000 words for the first half of the 18th century and found 290 POSS-Ing gerunds and only 2 that were nominal, one of which is the following from 1714

(57)   I do most solemnly declare,...that *my writing of these 'memoirs'* did not proceed from any desire of being an author...(Lockhart, *Memoirs of Scotland)*

While I have no doubt as to the accuracy regarding the development of the verbal POSS-ing gerund, it is unclear as to whether Fanego is claiming that there was full scale replacement of the nominal form by a verbal one.[63] While the style of (57) could be considered a bit archaic or stuffy, the use of a nominal or verbal form in these contexts is usually one of dialect and style and if Fanego takes her data to argue that these forms have disappeared altogether, it is a leap too far.

What did disappear (or rather fall into conventional disfavour) were some of the forms that began to verbalise in the 17th and 18th century,[64] such as Tajima's Type VI (e.g., this from 1673)

---

[60] Witness the various examples of coercion in the earlier semantics section.

[61] (Fanego, 2004) makes much of the need for semantic closeness, but does not elaborate further focussing more on the commonality of syntactic properties. It is here that I think that while (Fanego, 2004) dismisses (Houston, 1989) out of hand, there is room for both in the overall historical process.

[62] i.e. took a bare direct object.

[63] This is also the case with the determinerless Type III gerund. We have seen that it is alive and well in PDE, especially in presentation *there*-insertion contexts.

[64] Though one can find much earlier examples, this is their time in the sun, frequency-wise.

(58)   the noblest End is *the multiplying children* (Taylor, *Sermons*)

This type, along with some of the mixed verbal/nominal types types including:[65]

(59)   a. The *quickly doing of it* is the grace. (Ben Johnson, *The Alchemist, 1610*)
b. I began now seriously to reflect upon ... how justly I was overtaken by the Judgment of Heaven for *my wicked leaving my Father's House*, and abandoning my Duty (Defoe, 1719)
c. The Gentleman who writ this Play...appears to have been mis-led by an *unwary following the inimitable Shakespeare* (Steele, 1711)

Indeed, for a long period of history the gerund remained a hybrid form, whose mixed verbal and nominal properties allowed for examples such as the above. In Fanego's view it was both language internal matters as well as normative pressure that regularised the gerund and eliminated what was a very common form – the Type VI.[66]

Regardless of the specific details as to whether some forms were replaced completely by a verbal gerund or both a nominal and verbal form survive today, what is important is that in Fanego's chronology, the extension of verbal properties begins along a hierarchy of relative nominality. The first to verbalise were those without a determiner, the second the ancestor of the POSS-ing, and finally the rise and fall of the Type VI. As Fanego writes, "the more noun-phrase-like a sequence was, the slower it was to acquire verbal traits." (p.38)

The spread of the Type V to non-oblique positions

Thus far in history, the *ing* nominal has developed verbal properties, mostly in an oblique context. This is especially the case for the trajectory of the development of the PRO-Ing gerund. However, it did gradually develop external nominal syntax, that is it could appear in any position that a noun could, i.e. also in subject and object position. The regular use of PRO-Ing gerunds in subject and object position did not occur until the 19th century though a few examples are seen as early as the 16th century.[67]

Fanego traces the beginning of the spread with a change in some three-place[68]

---

[65]This includes the Type VI, as well as adjectives used with bare direct objects gerunds and other mixed forms.

[66]The instances of adjectives with otherwise verbal gerunds are far less frequent, and it is unclear as to whether only language internal factors lead to their downfall or that these were always marginal examples. As for the Type VI, the reader is directed to (Hindsill, 2001) for lengthy quotes from prescriptive grammarians condemning such 'illogical' uses as seen in (58).

[67]One of Fanego's samples of the 1500-1570 period found only 5% of of determinerless (verbal or nominal) gerunds in non-prepositional position. This contrasts with gerunds containing possessives or determiners that have a 30% non-prepositional use.

[68]Three place as the construction can be characterised as '**V sub. from X-ing**', where the

negative implicative verbs, all have a basic meaning of prevent or refrain. One of the earliest examples (with a nominal, determinerless gerund) comes from Wyclif (c.1400).[69]

(60)   lette men fro doing of iuil.
         prevent men from doing evil.

She then speculates that two place verbs, often with the same sort of meaning, expanded from taking on a purposive *to* infinitive to also taking a gerund. Compare the following, separated by two centuries.

(61)   a. Ye wol forbere now to do vengeance. (Chaucer, 1386)
         b. They come so to purpose, that hee can not refraine *telling them.* (T. Hoby, 1561)

The type of gerund use seen in (61-b) then spread to similar verbs, such as *avoid*, *decline* and *shun*, and eventually to broader classes of verbs, eventually replacing partly or completely the use of the *to* infinitive in emotive verbs, *remember* type verbs and verbs of intention.

    The use of the PRO-Ing gerund as a subject came last, an early example being the following

(62)   a profane idolatry of kings,...as if *being born of a certain race* could entitle any family to a right of violating with impunity all laws, both human and divine. (Walpole, 1778)

Fanego argues that the reason for such a slow spread had to with subjects being a special type of core complement.[70] In this view, there is a grammatical relations hierarchy, following along the lines of oblique, internal core complements (objects) and, finally external core complements (subjects), and that the spread of the PRO-Ing gerund followed the direction of this hierarchy, following (Timberlake, 1977)'s views on the directionality of actualisation.

Development of the ACC-ing gerund

The ACC-ing gerund is considered as a separate case of reanalysis, this time of a participial origin. Fanego notes that gerunds with accusative case never appeared in either pure nominal form or hybrid from.[71] For this reason, she views them as being unambiguously clauses. This then gives the suggestion that the ACC-Ing type did not result from the nominal gerund slowly taking on verbal properties, but a series of abrupt reanalyses of certain types of participial constructions.

---

preposition is considered a place.
   [69]Cited by (Visser, 1963-1973).
   [70]See (Valin & La Polla, 1997) for an in depth description of core and peripheral complements.
   [71]All examples of these involve true determiners or possessive pronouns.

However, the main culprit for the development of the ACC-Ing gerund is taken to be a certain type of participle, which precedes its main clause and controls the subject of the main clause.[72]

(63)　　a. And so *Vaughan's Testimonie being credited*, ∅ may be the material Cause of my Condemnation... (Throckmorton, 1554)
　　　　b. At last they found that which they expected, which was Sea-cole, *they following the veine of the Mine*, ∅ did dig forward still. (Taylor, *Pennyles Pilgrimage*, 1630)

It is a reanalysis of these sorts of constructions, which are preverbal, that explains the appearance of over 50% of the early ACC-Ing gerunds to appear in subject position, at a time when the other types of gerunds were still appearing mostly as objects of prepositions and to a lesser extent as objects. Note that the deleted subject in the main clause is only the subject of the participle. In contrast, the reanalysis is that the entire *ing* clause is the subject of the main clause. Moreover, these early examples of ACC-Ing as subject also display a factive reading of the gerund, which was far rarer among the early examples of PRO-ing in oblique position. Here is an early-ish example:

(64)　　*It not being the Time of the Day in which Company usually walk*, made the Place they were in extremely retir'd, and they had the good Fortune of meeting no Interruption. (Haywood, *The Fatal Secret*, 1725).

The subject of *made the Place* is not the expletive *it*, but that the time of day of walking is different.[73] Interestingly, what is not lacking is semantic closeness. The participle in sentence (63-b) clearly has an actional reading that moves the scene along the vein of the mine, while the examples given for the early ACC-Ing examples are clear factive. No explanation is given for this, and while Fanego does note that the factive reading (for gerunds) did emerge much later than the actional reading that it began with, she never explains how this meaning extension came about.

Unanswered questions

Keeping to the subject of participles, we will end the history section by looking at a phenomenon that is mentioned by neither Houston or Fanego. The contemporarily odd construction had a brief flowering in the 16th and 17th century.[74]

---

[72]The use of the emptyset by Fanego in the following examples is to indicate that the subject of the clause is identical to the subject of the participle that precedes it.

[73]I find the subject in (63-a) to be a bit ambiguous. Is it that the testimony was credited or the testimony itself that is clause for condemnation. Sentence (63-b) is much more clearly a participle.

[74]Examples from (Visser, 1963-1973).

(65)   a. The rosis yong, new *spreding of their knopsis.* (Dunbar, 1520)
b. Each one...Chas'd us away, till, *raising of more aid*, We came again to bind them. (Shakespeare, 1590)
c. Wenches sitt in the shade...*singing of ballads.* (1648, Letters of Dor. Osborne to W. Temple)

Visser notes that while instances appear as far back as the 14th century, it is not frequent until the 16th century and then falling into disuse by the 18th century. His suspicion is "that in the course of time the semantic difference between the purely verbal and mixed verbal + nominal forms in *ing* was so vaguely realised that the use of *of* after what was originally a participle...was no longer felt to be solecistic." (p.1204) As this time period coincides with the greatest amount of changes in the gerund,[75] Visser seems to be getting at this being an example of Croft's *Intraference*. At the time the semantic difference between appositive participles, the progressive[76] and the gerund were all rather close in that they all denoted actions, rather than something factive. While, Fanego is absolutely correct in her demands on Houston to have a syntactic (along with a discourse functional) mechanism for change,[77] it may nevertheless be the case that the various *ing* forms did have a sort of conspiratorial influence on each other. I find it rather hard to believe that the brief flowering of the examples in (65) are a mere coincidence, but may be a sort of blow-back phenomenon.

Indeed, it has been questioned as to how categorically the different *ing* forms can be differed in contemporary grammar,[78] and shall be explored in the next section. Concomitantly, the more gradient rather than categorial division is supported in that the supposed clear-cut division between verbal and nominal gerund still does not exist in ordinary speech, as hybrid forms still abound.

---

[75]Including very bizarre hybrid forms.
[76]A now obsolete form of the progressive, but common at the time.

(i)   in the meane time *I am building of A house* toe put my head in. (Calvert Papers, 1638)

[77]Interestingly, Houston, notes the similarity of discourse function between the use of appositive participles and the early oblique verbal(and nominal) gerund. It is closer than that. For example, the earliest verbal gerund example (*was bisi...in casting kirc wirkes.*) has a quasi-progressive reading, where the gerund certainly denotes an action. As factive readings did not arrive for some time, the semantic closeness between appositive participles and early verbal gerunds is much closer than she intimates.
[78]That is to say it ultimately may be hard to make a categorial distinction between verbal and nominal gerunds and participles. Deverbal nouns that denote a physical object such as *clothing* or *building* are a different matter altogether.

6.2.5   PDE and the Gradience that dares not speak its name

In both the account of gerunds in PDE (Section 1) and the history of its development (most of section 2), we have seen a strict grammatical distinction between (present) participles and gerunds, both of which belong to the verbal category. In contrast, the *Ing*-of nominals are considered to belong to the nominal category. It is only when examining Fanego, that an idea emerges of a gradient relation between the two categories, and while she uses this notion in order to account for the change in the gerund historically, it is seems to be the case that the split verb and noun in the gerund is not complete.

In both (Quirk, Greenbaum, & Svarvik, 1985) and (Huddleston, 2002b), there are intimations of this. While (Huddleston, 2002b) certainly makes a distinction between the *Ing*-of gerund and what we have been calling a verbal gerund, the distinction ends there. Instead, in a section aptly called 'The distinction between gerund and present participle can't be sustained' (p.82), non-deverbal *ing* forms are given a tripartite distinction between gerundial noun (*ing*-of), the gerund-participle form of verb (including the verbal gerund, present participle and its use in the progressive), and the participial adjective (an *entertaining* show). While granting that historically the present participle and gerund have different sources,[79] in PDE the forms are identical as far as internal syntax goes. That is when looking at modification by adverbs, the taking of determiners,[80] and plural inflection.[81]

(Quirk et al., 1985) goes a step farther. While they concur that there is no useful reason to distinguish between a gerund and a participle, they also describe a rather 'complex gradience' (p.1290-1292) from deverbal nouns to participles/gerunds with *Ing-of*.[82] Sadly, they do not go into too much detail other than a comprehensive analysis of some fourteen examples on this cline. However, the inference seems to be both in (sometimes) syntactic ambiguity, syntactic gerundial oddities, as well as a few novel (noun-like) interpretations of verbal

---

[79]Especially if we follow Fanego and view the primary source of the verbal gerund as changes emerging from the nominal gerund.

[80]But some doubts will follow soon in this matter.

[81]Only gerundial nouns can take plural inflection.

(i)   a. The killings of the birds must stop.
        b. *Killings the birds must stop.
        c. *Killings the birds, he stalked the meadow ferociously.
        d. *The nasty hunter is killings the birds.

This syntactic distinction between gerundial nouns and participles is the only one I have never seen exception to, either historically or in PDE.

[82](Quirk et al., 1985) use the more apt term verbal noun, as the *of*-marked object is optional. Verbal noun includes both *the horrific bombing of civilians* and *the horrific bombing.*

gerunds[83] suggesting that the verbal/nominal distinction is far from categorical. Indeed, taking a closer look at people's everyday speech shows that the 'verbal gerund' is not quite so restrictive when it comes to taking determiners. Moreover, while (Pullum, 1991) dismissed some of the more modern examples as deliberate archaisms,[84] the use of *the* with verbal forms remains alive and well.

First off, impressionistically, at least, it may be the case that cognitive verbs are most felicitous (when negated) in a hybrid form. Recall that modification by *not* rather than *non* or *no* is a hallmark of the verbal *ing* form. Nevertheless, they have little trouble in also being modified by *the*.

(66)   a. The worst is *the not knowing*. I think everybody holds out hope that their child will come back alive.[85]
b. No one knows for sure and it is *the not knowing* that has the Gonzalez family troubled.[86]
c. Congratulations! I know exactly how you feel: the shaking, *the not believing you're actually there*, the continental breakfast.[87]
d. But there are two ways in which I could sin here, firstly *the not believing it possible* at all under any circumstances, and secondly continuing to think ...[88]

With both *know* and *believe*,[89] people will often more explicitly say, *The not-believing bit* (or part, or thing).[90] Indeed this is the strategy which (Pullum, 1991) tries to explain away some of (Schachter, 1976)'s PDE examples. The examples are *Any talking loudly on your point will be punished* and *This burning the midnight oil has got to stop*. He says that they do need a special mention in the grammar of English, but only goes so far as to suggest that they should be compared to the 'hyphenated-compound-adjective' construction seen in phrases like *the easy-to-please image he has adopted*. This is a sort of the use of the Ing for as modification of a noun, which Pullum says are verb phrases used as *if* they were nouns denoting an activity. However, while he claims that Schacter's example represent non-productive constructions, those with *knowing* and *believing* are quite common.[91] Furthermore, the use of *the* with a verbal gerund/participle

---

[83]Which we shall look at in the next section.

[84]Examples from Dickens and Somerset Maugham.

[85]Taken from http://transcripts.cnn.com/TRANSCRIPTS/0202/27/ltm.13.html, a reporter being interviewed about a kidnapping in San Diego.

[86]http://www.boston.com/news/nation/articles/2005/10/16/a_street_of_lost_dreams_new_hopes, article in the Boston Globe.

[87]A commenter on a blog re: the nervousness of being in Seattle for an interview with microsoft. http://tejas.wordpress.com/2005/11/30/microsoft-interview

[88]www.anglicanrenewalministries-wales.org.uk/main/magazine/issue32/MNewsom.htm

[89]And much more generally.

[90]In the examples above, it is the case that some the hybrid nominals are a sort of list of activities that are parts of something else. These specific activities are the parts.

[91]Moreover, the more grammatically proper versions *the non-knowing* and *the non-believing*

is not restricted to these cognitive verbs, but is used much more widely. This even is the case with tense-modified gerunds and *being*.

(67)   a. Even the little original bits like Honey's fake illness and pregnancy aren't enough to cover the fact that he's done all this before, the romance, *the trying to impress the father in law, the being generally aimless*.[92]
b. But it isn't *the being nervous that causes the problems*, but the worrying about what being nervous might lead to.[93]
c. *The being nervous about it before hand*, and taking those laxatives was by far the worst part.
d. I'm done with *the feeling sorry for myself*, with the pity parties, with *the being sad all the time*.
e. In its defense, *the having been recorded live* (and not very well) to begin with, we probably couldn't have hoped for more.[94]
f. I'm working on the twirling, and *the catching it in the air*. Do you guys know SteveO, he's the master of broomtricks![95]

In all of these examples, it is certainly the case the hybrid nominal is some sort of activity, but these are not special or unique constructions that can be explained separately from gerund/participles in general, but rather need to be taken into account for any adequate syntactic description.[96] However, the notion of a verbal nominal being construed as an activity is not a special case of coercion, but rather the a standard interpretation of verbal nominals. Indeed, the 'act/fact' distinction discussed in (Abney, 1987) has to be modified. We shall see in the next section that the interpretation of verbal nominals is far more flexible than typically thought to be. But first, a final look at a PDE oddity involving modification by *no*.

6.2.6   A non-archaism

Another typical example in the literature of past-historical oddities is the following that has a verbal gerund with a determiner.

(68)   There is no enjoying life without thee. (Ben Johnson).

---

do also exist, but to my ears sound much worse.

[92]Someone on an *Eastenders* message board complaining about character development.

[93]The following examples found on various message boards and blogs.

[94]Interview with one of the people who worked on a live Monty Python DVD.

[95]A janitor in Los Angeles talking about what he does when he's bored. http://www.xanga.com/BohoRoohaha

[96]These are the times when I am quite glad not to be a syntactician. My experience is that if once one is aware of the supposed syntactic 'facts' of verbal nominalisation, these sorts of examples begin appearing everywhere.

(Pullum, 1991), when going through a list of oddities classifies this along with *my wicked leaving my father's house* and *constant handling the rod and gun* as a vestige of history.[97] (Denison, 1983) refers to a variant,[98] and groups it among 'rare' examples. (Abney, 1987), however, does acknowledge its use, but dismisses it from his analysis as it is an idiomatic construction. This seems no good reason for such as dismissal, especially as it is incredibly productive and common. To my knowledge, it is only (Quirk et al., 1985) who note the ubiquity of it. Furthermore, the determiner may also be *any.* Here are their examples.

(69) a. There's *no mistaking that voice.*
   b. There was *no lighting fireworks that day.*
   c. There isn't *any telling what they will do.*
   d. There must be *no standing beyond the yellow line.*

They do not mention it, but aspectual verbs also work rather nicely

(70) a. Once he gets a haircut, there'll be no keeping the girls away.
   b. There's no stopping us.[99]

This construction often has a modal interpretation. For example, sentence (69-b) could be said if there was a deontic proscription on lighting fireworks,[100] or that excessive wetness made it impossible to light fireworks – dynamic modality.[101]

 The deontic versions needn't even be in an existential *there*-construction, but a simple imperative:

(71) No hitting your sister in MY house, young man!

When we examine the more 'flexible' readings of verbal nominals, we shall see that the modal readings are not the result of participating in the negative existential construction, or an imperative, but are actually a quite common reading.

---

[97]Though he only says that these last two are 'shockingly ungrammatical'.

[98]*There's no making you serious a moment*(Sheridan, 1777).

[99]Common expression, and also the title of a Supremes song.

[100]To a Dutch audience, this may be hard to imagine. But coming from California, I can testify that it is the case there that the fire danger makes certain there generally is no lighting fireworks.

[101](Quirk et al., 1985) notes that generally the *Ing*-of nominal is best to express cases of no modality, such as *There was no shooting of prisoners*, which can have the interpretation that no one shot prisoners.

6.3 A second look: The semantics and mappings of verbal *Ing* forms

This penultimate section of the chapter looks at the interpretation of verbal gerunds.[102] For example, the some rather unexpected readings of verbal gerunds are to be shown – where 'manner' readings do exist, as well as the readings with predication of time. More broadly, it shall be shown that the 'act/fact' distinction does not exist at all. Rather, verbal gerunds have a an incredibly wide range of interpretations. Some of this flexibility, is already attested, as can be seen in what immediately follows. However, empirical evidence will demonstrate that even these broader viewpoints do not go far enough. After these are discussed, the sections leading into the formalisation, do contain some speculation as to the basic denotation common to all of them, and what could account for some of the rather surprising examples.

6.3.1 A revised typology

In a recent book, (Heyvaert, 2003) has a typology of what she calls 'nominal strategies adopted by gerundive nominalisations'. There is a cognitive grammar theory behind how the nominals are derived, that at most a passing mention shall be made. However, the resulting typology is quite useful, as with (Quirk et al., 1985) and others, she makes the distinction between action gerundives and factive gerundives. To summarise:[103]

- Generic reference, proper name strategy:[104] *Eating vegetables is healthy*
  This has an action reference, but is generic or habitual. This is the classic, third type of *ing* nominal.

- Individual situation, but proper name strategy. There is both an action type and a factive type:

  (72)   a. *Adjourning at 4:00* was impossible.
           b. I resent *being tagged a problem boy.*

- Individual situation, definite, specific common noun strategy. Both action and factive:

---

[102]At this point, it should be clear that the verbal gerund is a use of a more general verbal *ing*-form, for example, called the gerund-participle in (Huddleston, 2002b).

[103]A summary of her typology can be found in (Heyvaert, 2003, p.240).

[104]The notion of a 'proper name' and 'common noun' strategy originally comes from (Schachter, 1976). Simply put, determinerless nominals are said to function analogously to proper names, while those with a subject or determiner are said to function as common nouns.

(73)   a. his job involves *my answering the phone on his behalf.*
         b. *the country's making a statement is a lot bigger than partisan politics.*

- Individual situation, definite or indefinite, specific. Common noun strategy. Only action reading here.

(74)   a. *This burning the midnight oil of yours* has got to stop.
         b. *Some reporting from you* would be welcome.

The proper name vs. common noun distinction has to do with whether the gerund is subjectless or not. Definite or indefinite has to do with what sort of determiners can be used and the interactions with these that are found with ordinary common nouns depending on whether it is generic or not, count or not. Specific can be contrasted with generic, but still may be habitual. The major difference between specific and generic has to do whether it is a specific participant involved, rather than a generic use that applies to any participant who could be involved in the activity. While sentence (74-a) is specific in this sense, as it refers to a repeated activity that must stop, it is also habitual.

But perhaps what is most interesting here is that some of the specific, action nominals can (with a bit of rewording) be paraphrased as *that*-clauses.[105]

(75)   a. That we should adjourn at 4:00 was impossible.
         b. His job involves that I answer the phones...
         c. That you keep burning the midnight oil has got to stop.

None of the above can be paraphrased by 'the fact that', and cannot easily be interpreted as facts, but rather events or habitual activities that have occurred, will occur (or are occurring), or activities that are part of a larger activity. In contrast, the examples that Heyvaert labels as factive can be replaced with either a *that*-clause or *fact that*-clause, such as

(76)   a. I resent the fact that I am called a problem boy.
         b. I resent that I am called a problem boy.
         c. The fact that the country has made a statement is a lot bigger than partisan politics.
         d. That the country has made a statement is a lot bigger than partisan politics.

Thus far, it appears that gerunds can usually be paraphrased as a that-clause, but only sometimes by 'the fact that'. That is, both a gerund and that-clause

---

[105]The following examples also come from (Heyvaert, 2003). But, alas, many of her interesting ideas regarding factivity cannot be included here. The reader is encouraged to see Chapters 8 and 9 of her book.

can refer to an action/event or the proposition that it happened. It is only in the last instance that a 'fact that' paraphrase is allowed. But, when there is a factive reading in an oblique position something else happens.

(77)   a. We are pleased by *the fact that you work so hard.*
         b. *We are pleased by that you work so hard.
         c. We are pleased by your working so hard.

The last example shows that a garden-variety *that*-clause is not always factive. They often have such exclusively activity interpretations. Heyvaert calls the different readings nominalising strategies. In fact, there are a variety of other interpretations of 'verbal gerunds', including some that are very close to the 'eventive' readings, usually considered the province of nominal gerunds only. These shall be explored in the following sections. The following will show that gerunds have not only an action interpretation, but the action can be predicated of in certain ways not noted by the above typology – for example, an internal perspective on the action. Furthermore, there exist as well 'event' readings that should theoretically only be possible with nominal gerunds. Some of these are taken from (Quirk et al., 1985), while others are, to my knowledge, previously not mentioned in the literature. Importantly, few of the examples are idiosyncratic and exceptional, but rather perfectly normal uses in modern English.

6.3.2   Manner, reason and beyond: Toward a unified interpretation of verbal gerunds

While Section 1.2 gives the standard view that manner readings are the exclusive province of nominal gerunds, it is not always so restricted, as should be clear from the following examples. The first is from (Quirk et al., 1985, p.1291), and the second was overheard on BBC, Radio 4.

(78)   a. I dislike *Brown's painting his daughter.*
         b. Michael Vaughn's ingenuity and *his moving the fielders around* was light years ahead. (of Australian cricket captain Ricky Ponting)

For sentence (78-a), Quirk, et. al. claim that there are two readings for the sentence. The first is that the subject dislikes the fact that Brown paints his daughter. This much is uncontroversial. However, they give a second reading, which sounds suspiciously like a manner reading – 'I dislike...the *way* he does it.' (emphasis mine). (78-b) was overheard on Test Match Special, the BBC cricket program, and was said by ex-batsman, now summariser, Geoffrey Boycott, retrospectively comparing the captains of England and Australia in terms of performance during the last (2005) Ashes tests. The comparison is not that Vaughan moved his fielders around and Ponting didn't. It is rather that the way Vaughan did the activity was far superior to the way Ponting did it.

While it should be clear that verbal gerunds can have a process or action reading,[106] temporal duration can also be predicated of them.[107]

(79)   *Your driving a car to New York* took longer than I expected.

More generally, (Quirk et al., 1985) hold that verbal gerunds can have an action or process meaning (as well as a factive reading). If this is the case, there is certainly nothing odd about then predicating temporal duration of them, as processes and actions take time.

The existence of 'reason' nominals in English

To my knowledge, the final reading has never been cited for English, but is not unique in the pantheon of the world's languages. Reason nominals do exist (Comrie & Thompson, 1985, p. 357), and, given the proper matrix verb also exist for English.

(80)   a. I understand Mary('s) leaving town suddenly.
         b. I understand that Mary left town suddenly.

Sentence (80-b) is not a paraphrase of (80-a). Rather the first sentence has as a paraphrase

(81)   I understand why Mary left town suddenly.

While I have yet to see a semantics for *why* questions, and have no deep explanation for this phenomenon, it seems clear that some sort of modality is involved. We have also seen with the *There's no* construction' in section 2.6 that there are modal interpretations for verbal gerunds. Notice that with one or more modals, the 'reason why' reading can be gotten from a *that*-clause complement. Take the following sentences involving someone musing on being left out of a will.

(82)   a. I understand [no complementiser] he would leave me out of his will (but why leave you out?).
         b. I can understand that he would leave me out of his will.
         c. I understand why he left me...
         d. I understand his leaving me out...
         e. I can understand that he left me...
         f. I understand that he left me out of his will. (factive)[108]

---

[106]Rather than merely a *that*-clause, or a 'fact' that reading.

[107]From (Quirk et al., 1985) (p. 1064).

[108]Only the last has only a factive reading, but perhaps with a lot of intonation on 'ME', the 'reason why' reading can be obtained.

Processes in Inaction

(Lambalgen & Hamm, 2005, p. 222) correctly notes that "facts, results, propositions cannot be interrupted, but events or actions can", and goes on to formalise the notion of interruption for nominal gerunds (i.e. perfect nominals) only. While we certainly agree with the quote, the above discussion should convince one that verbal gerunds, when given a process interpretation, should also be able to be interrupted. This is indeed the case.

(83)   a. The bum, while dressed nicely, with a clear impression, *interrupted my taking pictures* and asked if I had a buck to spare for his bus fare.
         b. Yellow Cake interrupted *my watching 24 to talk to me.* I told her to call back after 24.[109]

Certainly, when verbal gerunds denote results they are able to be prevented, but it is also the case that activity denoting verbal gerunds can be prevented,[110] as the following examples demonstrate.

(84)   a. He claimed at the time that the injury prevented *his driving a patrol car*, using his gun or even working behind a desk.
         b. Beginning in 1947 Hayward suffered from a paralysis of his hands that precluded *his playing the piano*, but by 1950, he was able to play again.[111]

Steps toward a broader viewpoint

With these 'exceptional' readings in mind it is time for a wider look at the interpretation of verbal Ing forms, which shall be the subject of the remaining few sections. Moreover, if one accepts the categorisation of (Huddleston, 2002b) and (Quirk et al., 1985) in regard to complements of verbs of perception and recollection,[112] which are exclusively 'eventive', the need for a broader interpretation becomes clear.

   The following are yet more examples, all of which are perfectly fine for my ears that have non-factive readings, all POSS-ing. Some of them are quite astonishing, but in context are perfectly felicitous. Also, while the following examples have

---

[109]Example (a) from *voluntaryconfinement.blogspot.com/2005_11_01_voluntaryconfinement_archive.html*. Example (b) from *www.thebachelorblog.com/lexview.aspx?dart* = 2473.

[110]And allowed to resume after the prevention no longer is the case.

[111]Example (a) from *www.officer.com/article/article.jsp?siteSection* = 5&*id* = 27451. Example (b) from nfo.net/calendar/dec04.htm.

[112]E.g. *I remember travelling to New York* is the memory of the event, not merely a memory that it happened. For a comprehensive look at verbs of recollection, as well as both linguistic and philosophical implications, see (Higginbotham, 2003). Higginbotham considers the complements to verbs of recollection to be gerunds. In Chapter 4 it was shown that complements of perceptions are not gerunds, but 'aspectual phrases'. However, I think in the terms of (Huddleston, 2002b) these too would fall under the broader umbrella of gerund-participle.

actional or sometimes even manner readings, most of them are not paraphrasable as a *that*-clause either.

(85)   a. You worked out a nice *way* of telling the story – which was *you not looking at the camera.*
b. Still hard to say with my numbers since I've not fully broken in the car, have yet to have an oil change, and *my driving the car is pretty random.*
c. He told us that the bird feeder got smashed in the *process of his painting the house this past summer.*
d. *His building the ark, which took him so long a time*, and in which he employed so many hands, was a standing warning to them.[113]
e. He wondered suddenly if *his building the sewing cabinet* had been inspired by guilt.
f. Michio Kaku's ability and passion for physics resulted in *his building an atom smasher in his parents' garage when he was 17.*
g. Peter Vanderwal documents *his building an electric bike.*
h. Even if he's discussing *his building an invention*, his explanations are as engaging as they are educational.

None of the above are factive, and some seem coerced into a manner reading (e.g. (a) and (b)). (85-d) has a POSS-ing with a temporal restriction (but stated in a roundabout way). What becomes evident here is that there can be coercion when a gerund is embedded.

(86)   a. had decided to boycott me for *the mere fact of having written a book on Franco that was not a denunciation.*[114]
b. The act/process/way of his painting the house.

The first is, indeed, factive, but is an event that has been turned into something factive, while the second is an act or process.

As far as 'facts' are concerned, generally they are neither causative or caused. Facts can signify or mean something, be surprising or annoying, but as can be seen in (86-e) and (f), it wasn't facts that were inspired or caused, but the event itself. There is a subtle distinction, however. One could say that the fact that Michio Kaku built an atom smasher at 17 indicates that he was passionate about physics at a very young age.

As the above examples show, verbal gerunds can often be causative – which would restrict the meaning to events, actions, or emotional states, but certainly not facts. They also show that the default denotation of the verbal gerund is not

---

[113]This comes from a strange Christian website. The quote sounds archaic, but is just a bad attempt at a certain type of style.
[114]This is from (Heyvaert, 2003)'s corpus. I would say it is not the fact that caused the boycott, but rather the knowledge (and the concomitant political beliefs) that caused the boycott. The fact can justify the boycott however.

a fact or proposition, or even *that* something happened, but perhaps something closer to the activity or event itself.

Indeed, the specific interpretation of the gerund is quite flexible, and can be influenced by the discourse context or the linguistic context surrounding it. Taking Pullum's notion of modification of a noun seriously, it is clear that *Ing* forms have different readings when they are a prepositional object of a noun, taking on the noun's meaning. Restated now more generally, the following are only a few ways this can be done.

(87)   a. the fact of *Ving OBJ*.
          b. the act/process of *Ving OBJ*
          c. the way of *Ving OBJ* is to *Ving OBJ*
          d. the goal of *Ving OBJ*

As Pullum's recourse is to the adjectival modification version of the above[115] in order to account for certain hybrid nominals (Section 2.6), a return to history may show some justification.

(88)   The having been concerned in these practices abstractedly, was a circumstance which, according to his opinion... (Sir Walter Scott, *Guy Mannering*)

The hybrid nominal in (88) is actually a description of a circumstance. In many of the examples in (85), this is contextual from other parts of the discourse, and if it is not explicit, it can certainly be inferred. But, this is no different for ordinary nouns. The following is from someone commenting on their bi-polar disorder.

(89)   a. The depression part of it kept getting worse, particularly in the winter and spring.

Had they not used the word *part*, the reader would still infer that the depression is that part or aspect of the disorder that would worsen, depending on the season. This could just as easily have been written as *the being depressed part of it*. But facts are just as easily described in the same way. Staying in a historical mood, the following is a 1662 entry from Pepys' diary.

(90)   a. read the phrase about Mr C as that Sam thought Mr C *had* seen him and *this fact (the having been seen)* troubled him.

The last sections show that having a flexible interpretation of verbal gerunds is certainly necessarily, given the extant data. Next, we will see formalisation of some of the more 'heretical' readings. What is important to realise is that there needs to be no additional ontology. The interpretations acknowledged, facts, events, actions already exist in our formal ontology. What needs to be done is to

---

[115]Or as is common speech – *The touching my toes bit was the hardest part of the exercise.*

show how they can be constructed out of the basics.

Finally, it should be noted that this is not to say that there is no difference (in denotation or interpretation) between a verbal and nominal gerund. While we have seen both manner readings and temporal modification of verbal gerunds, a few things are not possible for verbal gerunds. These are plural inflection and such temporal predication as *happened at noon*.[116] I have no conclusive explanation for why this so, but it could be that nominal gerunds denote 'events' in the more ordinary language use of the word.[117] They are things that happen at specific times, and can happen repeatedly. They are also things that one speaks of conventionally as a generic event, i.e. *I hate going to a wedding/christening, etc.* For this reason, it could also be why it is normally only nominal gerunds can take such determiners as *every* or definite numbers such as *There were three shootings last night*.[118]

However, there is one rather exceptional construction with passive gerunds that defies any strict division between the interpretation of verbal gerunds as processes and of nominal gerunds as events. Superficially, at least, it appears that even verbal gerunds can have event readings. This is the subject of the following section.

What's happening here?

While the above section revises much of the standard interpretation of verbal gerunds given in section 1.2, a major restriction has remained. Perhaps connected with the inability of verbal gerunds to be pluralised, they also cannot be said to *occur* or *happen* in a specific time and space. Despite the existence of manner and action/activity readings for verbal gerunds to be fairly common, there remains the restriction on such predications as

(91) a. *Killing innocent civilians occurred/happened yesterday.
      b. *Innocent civilians being killed occurred/happened yesterday.
      c. The/A/Yet another killing of innocent civilians occurred/happened yesterday.

---

[116]But see section 3.1.1 for an interesting exception.

[117]A common English language expression is *That's a process, not an event.*

[118]However, *a lot of* is possible with verbal gerunds. This is an example of someone talking about continually watching a tape, in order to be believe that the university basketball team played so well. (www.cantonrep.com/index.php?Category=26&Reporter=Todd)

(i)   There was a lot of hitting stop, rewind, pause. I wanted to make sure I was seeing what I was seeing...

In ordinary language terms, this seems to quantify activities that are part of a greater event, rather than count events themselves.

However, it is not only possible, but fairly common to use a passive verbal gerund[119] with *happens*, but in a generic or habitual context.

(92)   a. *Civilians being killed happens in war.* It's unfortunate, and no wants to see it.
        b. *Being killed happens in an accident when a car veers off the road, being killed happens when a tornado slams into your house in the middle of the night.* Murder is different, murder happens when a person or persons gleefully take the life of an innocent person.[120]

Both of these examples are perfectly fine to my ears, and would be completely perplexing if not for them being habitual or in the second case, definitional. Nevertheless, the modification of temporal frequency of these passive verbal gerunds is also possible.

(93)   a. *Getting caught happens more regularly* than you might think, too. Professors read hundreds of papers, and they are attuned to the general level of writing...[121]
        b. Working on a submarine, *getting caught happens constantly.*[122]

But if the last two habitual readings are built up from a frequency of individual events, it is rather curious that the individual events cannot be said to happen using a passive verbal gerund. Presumably, solving this puzzle would take a greater understanding of the interaction of nominalised events with generic and habitual readings, and how the latter readings are constructed. The other matter of interest is that while I have found numerous such examples with the passive, I have found none using an active version. Given that the wide range of usage of verbal gerunds is rather astonishing, it is unwise to rule the possibility out completely, even though trying to formulate them sounds decidedly awful to my ears. The least anomalous possibility would involve a forced use in a definitional context like (92-b), such as a context where one is stating rules of a game.

Conclusion

This section has moved the denotational split between nominal and verbal gerunds much farther beyond an event/result differentiation or Abney's act/fact distinction. It is now clear that verbal gerunds have uses that have manner readings, causal powers, processes that can be interrupted or prevented, and be temporally

---

[119]Of the type seen in (91-b).
[120]First example from http://forums.thesmartmarks.com/, second from http://www.bittersweetme.net/blog/?p=1152.
[121]From www.easternecho.com/cgi-bin/story.cgi?4570, a newspaper story about the prevalence of plagiarism.
[122]From http://messageboard.tuckermax.com/archive/index.php/t-1551.html, about the difficulties of masturbating in close quarters.

predicated regarding the length of the process or activity. The one remaining distinction is that nominal gerunds can denote events in a way that they can be counted discretely and be specified as happening in certain time or place. Indeed, Vendler's terms of Perfect and Imperfect nominal are certainly prescient seen in this light.

In the final section of this chapter, some of the process and causal uses of verbal gerunds are formalised in the event calculus. Felicitously, while (Lambalgen & Hamm, 2005) does not recognise many of these possibilities, the denotation of verbal gerunds as a type of fluent (whose factive or result reading is further derived) means that the basic formalisation can remain untouched. What is needed then are ways of using the pre-existing nominalised fluents to construct scenarios that can account for phenomena demonstrated in this section, barring the still baffling use of a passive gerund that can co-exist with natural language *happens* predication.

6.4   Formalisation

Space does not permit a full explanation of the way nominalisation is formalised in the event calculus. For this the reader is referred to (Lambalgen & Hamm, 2005, Chapter 12). But, we shall give a sketch in what follows. There are two basic denotation types of nominals – perfect and imperfect, corresponding to nominal and verbal gerunds, respectively. A perfect nominal is denoted by an event type and an imperfect nominal denoted by a fluent. Thus, only perfect nominals are allowed to be arguments of the *Happens* predicate.[123] As imperfect nominals are fluents, they are allowed as arguments of the *HoldsAt* predicate.

The basic process of making VP denoting predicates is the same as the reification of predicates to produce terms of the language noted in the Interlude.[124] The perfect nominal is derived by suppressing the temporal parameter by existential quantification before turning it into an object.[125] For example, if the predicate $burn(x, y, t)$ means that x burns y at t, the perfect nominal *John's burning of the house* will be:

(94) $\exists t.burn(x, y, t)[j, h]$

which can then be an argument of the *Happens* predicate, i.e. *Happens*$(\exists t.burn(x, y, t)[j, h],s)$.

The imperfect nominal for *John's burning the house* is obtained by reifying the predicate as a fluent in the following manner.

---

[123]But recall section 3.1 for a natural language exception.

[124]Where e and the various f's are also products of nominalisation.

[125]Concrete scenario information can also be taken into account. See (Lambalgen & Hamm, 2005, p. 221 -224).

(95) $burn(x, y, t)[j, h, \hat{t}]$[126]

Importantly this fluent is not a proposition, but can be related to a proposition via the *HoldsAt* predicate.

$$HoldsAt(burn(x, y, t)[j, h, \hat{t}], s) \leftrightarrow burn(j, h, s)$$

That it isn't a proposition is quite important as formalisation of the deviant verbal gerunds does not need a change of ontology, merely constructions that use scenario information in a proper way to get process and causal readings.[127]

Finally, (Lambalgen & Hamm, 2005, Chapter 12) demonstrates in detail different constructions in which perfect nominals can be coerced into facts and results, and the past perfect for imperfect nominals[128] can be treated as facts or results.[129] With that said, we can move on to discussing some of the examples described in the previous sections.

6.4.1   Cause

Recall the previous examples

(96)   a. Michio Kaku's ability and passion for physics resulted in *his building an atom smasher* in his parents garage when he was 17.[130]
b. He wondered if his *building the sewing cabinet* has been inspired by guilt.
c. *His building an atom smasher* in his parents garage resulted in the vapourisation of their neighbourhood.

The first example describes an inner state that results in the building of an atom smasher. But as seen in Chapter 2, certain types of states can be seen as a gradual process. This is what is going on here. Imagine, that ability and passion gradually build up until Michio Kaku has no choice but to engage in dangerous experiments. This calls for a telic dynamics as one would find in an accomplishment. For example, *passion* could be an activity fluent driving a trajectory with the parameterised fluent *impulse*(x) which when the culminating

---

[126]We shall henceforth write $burn[j, h, \hat{t}]$ to abbreviate such terms.

[127]We are not formalising the manner readings and the use of 'happens' with verbal gerunds. My suspicion is that matters could be much more complex in these cases.

[128]Having won the race was gratifying.

[129]The intensional differences between the two types of nominals described in Section 1.2 are also covered.

[130]This is a rather hard to believe example, especially as Michio Kaku is now a theoretical physicist.

point is reached ($impulse(x)$), the imperfect nominal $build[mk, a, \hat{t}]$ is initiated, i.e. $Initiates(e, build[mk, a, \hat{t}], s)$.[131]

The reverse causal direction seen in (96-c) implies that once the atom smasher is completed, then a fluent *danger* is started that eventually culminates in the vapourisation of the neighbourhood. Again there is a dynamics, this time leading to an event *vapourise*. The relation between the imperfect nominal and the fluent *danger* can be represented as

$HoldsAt(build[mk, a, \hat{t}], s) \rightarrow Initiates(e, danger, s)$, where e is the finish event taken from the scenario of *build an atom smasher*.

6.4.2   Processes and Temporal Duration

Observe the following examples

(97)   a. Taking the train to Nijmegen took three hours because of the NS' incompetence.
b. John's driving the car to New York took longer than expected.

The temporal predication is a bit subtle. The time profiled is not just the time profiled by the basic activity fluent of the scenario, but rather the entire 'event'. Example (97-a) has happened to me on more than one occasion. The 'three' hours mentioned is not merely time spent on the time, but encompassed to include the time waiting in the Central Station for the delayed train, and being kicked out of the train in Arnhem due to some problem or other and having to wait for another train to reach the final distance. Similarly, John's driving the car took may have taken longer than expected because of mechanical problems and having to get the car fixed, etc. For longs parts of the journey, no driving is going on at all, yet it is still conceived broadly as part of the *driving the car to New York.*

The question then is how to derive a temporal profile. The suggestion in (Lambalgen & Hamm, 2005, Chapter 12) for obtaining the temporal profile of *Deborah's singing a Poultry's tale* is to use the parameterised fluent $p(y)$ and define a fluent *increasing* that captures the time-points when the parameterised fluent is increasing. As can be seen in the above discussion, using this technique to define a fluent based on increasing distance will not do. In the context given for the above examples, distance covered remains unchanged for long periods of time, and this is why the trips take such a long time. In fact, the only usable measure is in fact, the time between beginning the journey and reaching the destination. It is the time in between these points, whether one is moving or at the mechanics, that profiles the journey.

---

[131]By using the imperfect nominal rather than an ordinary *building* fluent that is part of the scenario, the atom smasher is built, not just started.

Assume the scenario for *John's driving the car to New York* includes two punctual events – that of *leave* and *arrive*. For the process, we then create a nominalised fluent derived from $(drive,j,c)$, i.e. $drive[j,c,\hat{t}]$. From these[132] we can profile the time. For this fluents $clock_d$ and $time_d(x)$ are needed. Intuitively, the clock begins once the journey is started and terminated when it ends. Once the fluent $clock_d$ starts ticking, it drives the fluent $time_d(x)$ by a linear dynamics. The following scenario then emerges:

---

[132]Using a variation of a scenario in (Lambalgen & Hamm, 2005, p. 92).

1. $Initially(time_d(0))$

2. $Releases(leave,time_d(0),s)$

3. $Initiates(leave,drive[j,c,\hat{t}],s)$

4. $Initiates(leave,clock_d,s)$

5. $HoldsAt(time_d(x),s) \rightarrow Trajectory(clock_d,s,time_d(x+d),d)$

6. $Terminates(arrive,drive[j,c,\hat{t}],s)$

7. $Terminates(arrive,clock_d,s)$

Thus, the *leaving* and *arrival* events profile the time of the journey, and this information can be used to say that the time taken was, e.g. longer than expected.

6.4.3   Interruptions and Preventions

Recall the examples in (84), repeated here

(98)   a. He claimed at the time that the injury prevented *his driving a patrol car*, using his gun or even working behind a desk.
b. The bum, while dressed nicely, with a clear impression, *interrupted my taking pictures* and asked if I had a buck to spare for his bus fare.

Sentence (98-a) can be interpreted to mean that once the injury happened, the nominalised fluent of *his driving a patrol car* cannot hold. If we let *injury* be the event that represents the injury and $drive[h,car,\hat{t}]$ be the imperfect nominal corresponding to the verbal gerund, (98-b) can be captured by the following integrity constraints:

$?Happens(injury,R), R < now,$ `succeeds`, $?HoldsAt(drive[h,car,\hat{t}],s), s > R,$
`fails`

Sentence (98-b) describes an event of a homeless person interrupting someone in the middle of an activity – the imperfect nominal, *my taking pictures*. The interruption event is quite simple to characterise, it is merely a terminating event of a fluent, in this case $take[m,p,\hat{t}]$. The relevant addition to the scenario is an event *beg* that denotes the asking for money, which stops the taking pictures activity. It can then be resumed by a (re)initiating event, which continues the picture-taking from where it has left off.[133] Thus, the following is added to an existing, picture taking scenario:

---

[133]Although nothing is said about this possibility in the above example.

$Terminates(beg,take[m,p,\hat{\mathrm{t}}],t)$

which is the interrupting event. The dynamics of the scenario ensure that the picture taking continues where it was temporarily abandoned. For simplicity, and as the sentence makes no explicit mention of a resumption, the integrity constraint for sentence (98-b) is

$?Happens(beg, R), R < now$, `succeeds`,

meaning that after R any picture-taking has stopped. But it is possible that a *resume* event can come at a point after R, allowing the activity to resume.

6.4.4   There's no ...

In Section 2.6 we saw a peculiar construction with examples like:

(99)   a. There's *no mistaking that voice.*
          b. There was *no lighting fireworks that day.*
          c. There isn't *any telling what they will do.*
          d. There must be *no standing beyond the yellow line.*
          e. There's no pleasing some people.

There are numerous senses to this construction, including different types of modal senses, most of this that are beyond the scope here. However, there is one meaning that can be captured. A way of looking at example (99-e),[134] is that no matter what is done, some people are not pleased (by it). In terms of the Event Calculus, it is simply that no matter what event happens, the imperfect nominal denoting by the verbal gerund does not hold. Let *pleasing* be the imperfect nominal for *pleasing some people*, essentially the state of 'being pleased'. An integrity constraint is then added to the scenario that states:

$?Happens(e, t)$, `succeeds`, $Initiates(e,pleasing,t)$ `fails`

---

[134]Ignoring the generic part of it.

Chapter 7

Endnote

7.1 Conclusion so far

Before closing with some concluding thoughts, a brief summary of what has come before is necessary. Chapters 2 and 3 are foundational, and thus necessary for understanding the following three chapters. The Interlude (along with the appendix) shows more precisely how this can be done. This includes a computational system that shows how, inferentially, the flexible nature of aktionsart can be treated properly, and also accounts for the complications of tense and aspect comprehension discussed in the second half of Chapter 3.

Chapter 4 on perception verbs and their complements uses ideas discussed in the earlier chapters to give a comprehensive semantic account of their workings. After a discussion of perceptual reports' logical properties, concluding that they are a strong default, rather than a strict implication, the discussion moved on to the syntactic and aspectual properties of the complement. Here, the discussion in the earlier two chapters comes to the fore, as a good deal of empirical data is brought out to show that many of the supposed syntactic constraints advocated by (Felser, 1999) really lie in the realms of interpretation, such as the possibility of stative verbs being used in the complement. Ironically, one of Felser's major contributions – the revival of the Progressive Hypothesis – had been rejected for decades due to an over-reliance upon standard English to determine syntactic structure. (Akmajian, 1977) discusses just such a possibility and then rejects it on the grounds that corresponding sentences in the progressive are ungrammatical. For example on p.446 he notes that while PVC's with *needing, weighing* or *knowing* are allowed as complements of *imagine* (and equivalents) or *see* in the imagination sense, there are not allowed to appear in the progressive. Thus the PVC *Ing* complement cannot be a true progressive:

(1)   a. John is pretty hard to picture weighing 300 lbs.
         b. I just can't picture John knowing the answer.

    c. I just can't see myself needing any more drugs.
    d. *John is weighing 300 lbs.
    e. *John is knowing the answer.
    f. *I am needing more drugs.

Enough has been said about the status of sentences like (1-e), but it is also the case both *need* and even *weigh* can be felicitous in the progressive. Since both *need* and *want* are incredibly common in the progressive, I am quite mystified at Akmajian's assertion. *Weigh* is a rather special case, but is, for example, used by boxers in training, a context where a small variation of weight can mean being able to fight or not.

(2)   I am weighing 155, so I am going to have to weigh in with rocks in my pockets, but I am going to do it.[1]

Next, empirical evidence is brought in to show the claim that the temporal simultaneity between the matrix perception and complement cannot be syntactic, at least in the form of an aspectual phrase as it stands presently. Instead, a move is made to a semantic treatment of tense, allowing varying temporal interpretations of complements of aspectual verbs and verbs of intention, with a suggestion that the variation parallels the use of a 'real' present progressive and the futurate version.

    That done, the final piece of the puzzle is to determine the aktionsarten associated with *see* and *watch*. While *see* has been noticed to be anomalous by previous authors (e.g. the quote from (Huddleston, 2002b)), to my knowledge, it has never been worked out in depth. Using the techniques exhibited in Chapter 2, this is successfully done, showing a wide variety of event types for *see*, and exposing a rather unexpected interaction between a progressive *watch* and a simple aspect complementing, that while making sense semantically, is still slightly baffling.

    Finally, the logical properties and event-type, tense and aspect variants are all brought together in the final section of the chapter – the formalisation. There it is shown how the Event Calculus can be extended to account for complements giving rise to multiple event semantic structures. All of the varying aktionsarten are developed in differing constructions, with the predictions made as to the temporal profiles and implications emerging as desired. Furthermore, both with *watch* and certain event pairs such as *steal/borrow*, we have managed to account for extensionally equivalent, but intensionally different events, reinforcing the point of (Zacks & Tversky, 2001) about the importance of intention to a definition of event.

    Chapter 5 is an exploration of causatives, with the culmination being an in-

---

[1] Interview with Oscar de la Hoya from http://www.fightnews.com/trillo56.htm. This use of *weigh* is also common among fanatical dieters and body-builders.

tegration of various types of causative types into the Event Calculus. While the semantics literature has certainly noted that lexical English causatives cannot be reduced to one type of event structure (e.g. (Levin, 2000)), it is only by examining literature in the typological field that the full range of possibilities of causation types and participant roles becomes clear. While English tends not to have the same range of syntactic and morphological coding options as some languages, all of the semantic types and variations are shown to exist, if only by integration in context.

Continuing the investigation into English causatives, it was shown that both the causative alternation is actually more flexible than previously realised and that construal plays a large role in showing that certain lexical causatives, resultatives and 'way'-constructions with certain verbs or participants can be perfectly felicitous. Concomitantly, the brief look at the diachrony of a few typical causative verbs shows how the notion of participant and causation type can widen over time in a well-known pragmatic extension process, giving any synchronic, strict grammatical judgment (based on semantic grounds) a weak foundation.

Finally, the various causation types and participant roles were investigated in their various tense and aspectual guises, working out a variety of different temporal profiles and implicational paradigms that are associated with different forms of causation. Formally, different scenario templates were developed, that successfully accounted for the above, giving a formal analogue to such notions as direct and indirect causation along with the participant variants of manipulative and directive causation. Moreover, this continues the work of the formal section of Chapter 4 by extending the Event Calculus formalism to the coordination of multiple events, and the various temporal and telicity variations that exist. Showing formally the full range of possibilities reinforces the notion that construal, more than any other factor, plays a dominant role in acknowledging a loosening of restrictions restriction in a causative's natural language expression. That is, many example deemed ungrammatical in the literature were found to have numerous naturally occurring examples. While they may not be acceptable to all speakers of English, they certainly show that, in context, certain perspectives allow heretofore ungrammatical utterances to be utterly felicitous. Why certain speakers (and authors) do not realise the full realm of possibility is touched upon briefly in the next section.

Chapter 6 on English *ing* nominalisation returns to complexity of *ing*-forms, first touched upon in Chapter 4. However, the major contributions are mostly of an empirical nature. After examining the 'common knowledge' of the syntactic and semantic properties of nominalisation, the history and development of the various constructions are examined. Taking the notion used by (Fanego, 2004) seriously, of a non-categorial class membership to account for the development of the gerund, it is then shown than even synchronically it is premature to promote a strict differentiation between verbal and nominal structures. Indeed many of the examples that confound any strict categorisation cannot be considered 'one-offs'

or mistakes as often a general and common patterning is demonstrated.

In parallel, the interpretation of verbal gerunds (or *ing*-participles if that is preferred) is seen to lie in the same murky territory. While it is established that verbal gerunds should be interpreted first as eventualities, rather a derivative result or factive reading. This contrasts with an ordinary language notion of event, that can be said to be the default interpretation of a nominal gerund. Nevertheless, while verbal gerunds cannot be predicated to happen at a specific place and time or treated as individuals via quantification, many other readings that previously thought impossible turn out to be rather common. Temporal duration, manner and even generic/habitual happening can easily be predicated of them.

As their default denotation is as a fluent in the event calculus, the eventuality interpretation has already been constructed. As an extension to the scenarios found in (Lambalgen & Hamm, 2005, Chapter 12), many of the readings such as temporal duration and interruption are integrated into the formal system. However, not all interpretations found (such as 'happens' with a passive gerund) are at this point completely understood. Indeed, while there is a degree of speculation on this construction as well as others such as the previously unattested existence of reason nominals in English, it is clear that there is still much work to be done on the matter.

7.2   Found in the stars

One of the major themes that runs throughout this thesis is that there needs to be a perspectival reversal when it comes to distinctions of ungrammaticality, especially when it is (lexical) semantics that is at issue. Too often, what is captured is only the default reading, with deviant readings ruled out by the grammatical theory in question. Worse yet, a grammatical distinction can sometimes be little more than that of a distinction in dialect. This is a point that doesn't only refer to the grammaticality status of sentences where a semantic shift is ignored. Often, arguments concerned with morpho-syntax and word-formation to support the poverty of the stimulus argument are grounded upon a shaky empiricism. (Pullum & Scholz, 2002) have numerous examples where a supposedly ungrammatical construction that is meant to illuminate certain facts about our innate knowledge of language is little more than conventional dialectical difference.[2] For example they examine a claim made on supposedly innate constraints on noun-noun plurals. The proposed constraints do correctly describe the standard American dialect; however, the standard British dialect does the opposite for some compounds (typical example: *drug problem vs. drugs problem*). However, this is certainly not to say that there should be an 'anything goes' approach to

---

[2]This is not in any way to say that the reasons for these differences (be they syntactic or semantic) are of no interest or do not need to be explained in a systematic way.

language.[3] Nevertheless, the methodological point is that when one is trying to make a point regarding a *universal* constraint based upon, say, one particular accident of a standard English dialect, one needs to make sure that the grammatical constraint in question is indeed 'universally' or 'innately' ungrammatical.[4]

Despite the appearance of a wealth of 'deviant' actual examples that are normally ruled out for reasons of a semantic mismatch, I would also not want to claim that anything goes in lexical meaning.[5] Were this the case, not only would there be no particular purpose in studying language, but communication itself would be a rather random and hopeless task. Instead, the view advocated is that there is a great deal of structure and systematicity in semantics, and these anomalous examples only serve to support this notion. and are used to discover where the structure is and how it works. In Chapter 3, while it was shown that stereotypical scripts on events will often have a cultural side, world knowledge, causal knowledge and conceptions of time do go a long way in restricting our construals, and they way we build up structure for specific events.

As noted in Chapter 2, it is not the case that any eventuality type cannot be coerced into another one. Perhaps the clearest example is to expand on some thoughts of (Comrie, 1976a). While the verb *cough* can be used[6] in an activity construction, the interpretation will always be of a series of coughs, not a single one. Yet in 'objective reality', a single cough takes time and has some structure such as a beginning, middle and end. But on the human scale, we perceive such an event as instantaneous, and it is hard to imagine situations where an activity view of a single cough could be conventionalised as an activity or accomplishment. Were our perceptual system to work something like a super slow motion camera, conceptions could be very different.[7]

That said, a great deal is still possible. The notion of intention and its importance has been given a great deal of prominence. Normally, it is animate agents that have intention. But it is also the case that humans are also gifted at imbuing intentions to anything. Stories and myths are a prime example of this, but it can also emerge in ordinary language use.

To repeat an example from Chapter 6, it is common in the literature that

---

[3]This applies to event structure and perspectival shifts that make up the bulk of this thesis, and syntax even more so. I have seen no syntactic theory of any theoretical persuasion that does not involve a good deal of structure.

[4]Whatever this might mean.

[5]The same claim would be made for morpho-syntax as well. Despite my scepticism of some of the supposed examples that are said to support one flavour or another of syntactic theory, it would be absurd to claim that there is no systematicity and structure in language use. Leaving all cognitive issues aside, from a purely communicative perspective, an unsystematic language could never be successful.

[6]In some languages.

[7]It could be the case that in a class on the respiratory system, such a construal could be made, with the help of modern video technology. But it does seem inconceivable that it could be a conventional part of language use.

such sentences such as

(3)   *The river froze itself (solid).

are ungrammatical. Certainly, looked at out of context, sentence (3) really makes
no sense. But an examination of the family of causative and resultative con-
structions in English show that (with a transitive) the subject is a causal agent
in the process.[8] While rivers can normally be causal agents of flooding, decay
or humidity, simple world knowledge tells us that they are normally victims of
freezing and not a cause of it. Nevertheless, examples were found from science
fiction stories[9] where rivers can indeed freeze themselves. So then the methodol-
ogy is reversed from banning certain examples as ungrammatical to defining the
meaning of the construction. Sentence (3) is then not ruled to be ungrammatical,
but predicted to have a certain meaning. The oddity (and common attribution
of ungrammaticality) then resolves to whether or not this meaning is conceivable.
Normally, due to the nature of the world, rivers and soda bottles do not have the
power to freeze themselves. But to repeat an example from the Chapter 5:

(4)   a bottle of country club strawberry soda froze itself and exploded in my
        fridge tonight in a violent act of protest.

But given the right context, a different world physics in the river case, or the
attribution of agency in the soda bottle example (what is more intentional and
agentive than a act of protest?), the stars can be removed, and the constraints
are then interpretative rather than expressions of impossibility.

    In the Preamble, we suggested that a goal of the semanticist should be to
predict what kind of meaning a construction should have, if it is felicitous, rather
than being too inflexible and predict them impossible. The formalisations pre-
sented in various sections are examples of this. By taking coercion seriously in
a formal way and allowing shifts between aktionsarten and changes brought in
by context, we show how these apparent semantic oddities can be accounted for
in a systematic way. The new event-type and aspectual profile will then have a
predictable meaning.[10] Of course, the mere possibility of an instance of coercion
(formally) is not enough to guarantee that a speaker will find its natural language
instance felicitous. This is discussed in the following sections.

7.2.1   Construal and perspective

Given that it is now clear that there is certainly more flexibility in usage than is
often admitted to, the question then arises as to how this comes about. In the

---

[8]In a transitive construction.

[9]Where we can suspend our disbelief regarding ordinary causality.

[10]There is also the computational machinery meant to account for tense and aspect interpre-
tation.

cases that are the domain of this thesis, the answer is: construal. At the end of the last section, we were reminded that in the proper context a soda bottle can be said to freeze itself. This is because the speaker is construing (for discourse purposes and not making a metaphysical claim, presumably) that the soda bottle is a causative agent. Causative agents can freeze things including themselves, thus the construal allows its use in an otherwise banned causal construction. In general, it is probably the human capacity for perspectival shifts that brings about different ways of conceiving the same 'event'. This could involve a finer or coarser grained view of specific spatio-temporal phenomena, attributing agency in non-standard ways, or taking additional 'cultural' facts, that come up in the stealing/borrowing examples.

So far these are very specific and rather idiosyncratic examples, and it is perhaps more important that these are examples of more general semantic phenomena where perspective and construal can lead to violations of a language's defaults. In the first chapter, we saw not only examples of coercion in English tense-aspect constructions, but also that certain verb classes may be conceived as states in one language, but activities in another. For example,[11] in English, *dry* is typically either an inherent (*Death Valley is dry*) or transitory state (*The shirt is dry*), where a bit of rainfall in the former is does not make it untrue, but would the latter. While English also has a causative use of *dry* (*The shirt is drying*), (Foley, 2005) points out that in some Australasian languages whose speakers live in a rain-forest, *dry* is never an inherent state, but rather the consequence of an activity.

(Croft, 2000b) makes the distinction between *aspectual type* and *aspectual class.* Aspectual types are the universal semantic structures[12] that we have seen in various forms, both semantically ((Moens & Steedman, 1988), (Lambalgen & Hamm, 2005), etc.) and psychologically in the work of (Zacks & Tversky, 2001).[13] Aspectual classes are language specific and there is no direct universal mapping between a verb in a particular language and an aspectual type.

Moreover, even within a language a particular verb can map to more than one aspectual type depending upon context, and the construction that the verb appears in. (Croft, 2000b) labels what we have been calling coercion as construal operations, i.e. 'a reconceptualisation of the situation's aspectual structure.'(p. 4) There are two basic types – the first is when there is simply a shift on what part of the event is profiled (inchoatives, run-up achievements, etc.), but there is also the more substantial shift that 'involves a more substantial restructuring of the aspectual contour of the situation, using a combination of encyclopedic knowledge

---

[11]The transitory/inherent distinction is taken from (Croft, 2000b). To remind the reader of a distinction made in the first chapter, this is a sort of stative analogue of the single instance/habitual distinction that can be made for activities and accomplishments.

[12]These are more commonly called event-types.

[13]Various authors differ in their typology and granularity. For example, (Croft, 2000b) cites 18 different aspectual types.

about the situation, context-specific knowledge of that particular instance...and general cognitive operations such as shift of attention.'(p. 4)

That is, a default mapping can be made between aspectual class and aspectual types. Generally, verbs such as *sit, know, hope* are classified as statives. But it is better to look at this class as having prototypical statives such as *know*, where conventionally there is no activity construal allowed and more peripheral cases such as *sit*, where both stative and activity uses coexist happily in standard English. Nevertheless, we have found examples where *know* can be used in activity sense[14], showing that construal from state to activity is even possible with a core member of the stative aspectual class.

Lurking behind this is an idea of some sort of way of measuring what might be seen as, say, a prototypical or peripheral stative.[15] I have no simple formula to do this, but certainly a variety of factors contribute to this categorisation. Doubtless the way humans perceive the world and specific environmental factors (as seen with *dry*) will have a role to play. Moreover there is no reason why in a particular language group particular concepts are not conventionalised to be one type of eventuality rather than another. Ultimately, one person's intuitions are not enough, instead there would be a need to do a mass of empirical studies of speaker's judgements about characteristics of particular events (and contexts to play with). Even then, there is little doubt in my mind that even within a particular linguistic and cultural group, response would differ dependent upon the normal sociolinguistic factors of class, education, age, sex, etc. However, we have seen enough to know that while our shared human abilities and shared world will give rise to a great deal of cross-linguistic commonality, this sort of classification will not be universal.

7.2.2   Impact

In fact, taking the possibility of a construal shift from state to activity seriously can lead to a clarified view of supposed problems. As seen in Chapter 4, it is often held that a lexical stative[16] complement is ungrammatical in a perception verb construction.[17] As demonstrated, this is absolutely not the case. But what is the case is that in this type of perception verb construction the complement is interpreted as an activity.

Making the distinction between aspectual type and aspectual class causes this

---

[14]And indeed in some dialects it may be as conventional as *sit*, which is a paradigm of an English verb that straddles the state/activity border.

[15]Alas, these concepts will have to remained vague and undefined. However, think of it along the lines of class membership discussed in Chapter 3.

[16]Remember, stative here means a member of the stative aspectual class.

[17]A recent volume (31) of *Theoretical Linguistics* explores the issue of the Davidsonian event variable and this 'ungrammaticality' is a major issue.

debate to vanish.[18] The constraint should be that the aspectual *type* of the complement cannot be a state. However, as we have seen that in the progressive construction, verbs that are normally classified in the stative aspectual class can be construed as activities, and must be if they are used in the progressive construction. Similarly, when the message board poster said that he saw a particular talk show host 'be intelligent',[19] he meant that the host's behaviour was indicative of intelligence (and he saw that). Like it or not, Bill Clinton can feel our pain, but this is more an indication of him being empathic enough to infer this from someone's behaviour, not because of the emotional equivalent of mind-reading.

As already noted, by taking seriously construal operations and the meaning that emerges from combining a lexical item (that has a default aspectual class) with the construction it appears in, a different view on what the goal of (event) semantics emerges. This is now clear for many of the examples in this thesis, especially those of causative/resultative constructions and the many different examples of coercion seen in Chapter 2.

The analysis of the usage of nominalisation is much more complex, as so much variance has been uncovered. Certainly, it has been shown that the simple Act/Fact division between nominal and verbal gerund is completely inadequate. While there are certainly contexts that do require a nominal gerund, they are more infrequent than often is thought to be the case. Nevertheless, it is certainly possible to have even manner readings for verbal gerunds. My guess is that this is possible because of construal, of course in a more broad sense than aspectual shift. Take Geoffrey Boycotts comment about the England cricket captain's 'moving the fielders around' better than the Australian's. I think what is key to this case is that we have already seen that even as a default, verbal gerunds do have an 'action' denotation ((Heyvaert, 2003), (Quirk et al., 1985)), and not merely a factive interpretation. Once this is allowed for, then it is no great leap for a speaker then to be able to use a verbal gerund in a context of evaluating the action. In this case, this is not in reference to one event (conventionally speaking) but rather a habit of behaviour over two months. The example that includes 'my driving the car is random' (to mean the manner and frequency of driving the car) example works upon the same lines.[20]

---

[18]Not the debate over Davidsonian event variables, but the usefulness of the supposedly ungrammatical examples.

[19](Higginbotham, 2005) agrees that 'states' can be used in perception construction when construed as activities he claims that for unrelated syntactic reasons, the copula is banned. Thus 'I saw Mary happy' is allowed, but not 'I saw Mary be happy.' This may be a dialectical variation, as there are plenty of examples of 'be' with perception constructions and in general with small clauses 'He made me be happy.

[20]In general, the fact that the syntactic differences between a nominal and verbal gerund are not quite as clear cut as often assumed may also have something to do with the more complex mapping.

7.2.3   Convention

Of course, this is not to say that any of the above examples will be accepted (or used) by a majority of English speakers. Certainly the use of gerunds is a very complex process that varies over dialect, register, style, and the linguistic material that surrounds it.[21] That Geoffrey Boycott comes from a part of England,[22] where *ing* is used in many more contexts than, say, standard American or British English probably has a role to play.

While this is not the place to explore the general impact of convention on language and language change, we can restrict our focus to its interaction with lexical aspect. (Croft & Cruse, 2003, p. 102) identify two kinds of convention involved. The first is the language specific mapping between verbs and aspectual class, which will (in the particular language) constrain the possibilities of construal.[23] But within an aspectual class some exemplars are more flexible than others. This can be because of the nature of reality or human cognitive capacities.[24] But once an novel interpretation is made and expressed it, the construal made can be viewed as natural and be made a default. For example, (Croft, 2000b) cites the example, *She is being tall* as a non-conventionalised construal (in a situation where a girl gets up on a box to be taller than her friends). This is a fanciful example, though interpretable, that would not be accepted by many English speakers. The same type (state —> activity) coercion is seen in *She is being silly*, but this example (unlike the former) would probably be accepted by the majority of English speakers and is very commonly used. My guess is that *be tall* is much more of a core, inherent state, that the subject cannot turn on and off at will.[25] With *be silly* this is much easier seen as a temporary state that can come and go quite easily, making the construal much easier, and most likely more frequently used. Of course, it could be the case that in the future, a sort of fashion comes in where one wears platform shoes to signify sexual availability. The expression *He is being tall, tonight* could then be conventionalised to not only express a temporary state of tallness, but also to mean that someone is out 'on the pull'.

---

[21]Context and argument structure may also play a role. For example, an opinion piece by Norman Solomon is entitled 'How long will Moveon.org fail to oppose bombing Iran?' (http://www.alternet.org/columnists/story/35072/). While *Bombing Iran* is a 'Pro-Ing' nominal, context tells that the agent of Bombing is not the subject of the interrogative (an anti-Iraq war online activist group), but the U.S. government. Were confusion a possibility, then using *the bombing of Iran* would be used to separate any possible link between the matrix subject and the agent of *bombing*.

[22]He is from Yorkshire, and unlike the majority of his colleagues is not middle class.

[23]For example, the language have the capacity to allow shifts to be made between different aspectual classes.

[24]In (Croft & Cruse, 2003), these two are other constraints that are involved in the construal of interpretations, along withe ubiquitous contextual constraints.

[25]Though viewed over a longer temporal perspective, *getting taller* is a conventionalised construal.

Moreover, some examples of conventionalised construal are of such a default status that they hardly seem like coercion at all. For example, *drink wine* is an atelic activity and *drink a glass of wine* can[26] be a telic activity. This observation is stating the obvious, but this too is an example of coercion,[27] but so common and ingrained in the language that calling it 'coercion' seems to do an injustice to its normality. But it is coercion, nonetheless.[28] Finally, in the previous chapter, we saw how the argument structure and participant conditions of various causatives have evolved over time. *Kill* was initially quite restrictive, but has evolved to be conventionally applied in a variety of direct and indirect contexts. Others, such as *disappear*, while used by some speakers as a transitive causative[29] is undoubtedly still strange and unacceptable for many speakers.[30]

Bringing matters back to the beginning, the example of *is knowing* is a prime example of the interaction of construal and convention. In standard English, it is certainly unconventional, but can be used, whereas it seems that it certain dialects it is a rather common construction. This difference and acceptance of its usage is something best answered by socio-linguistics, but I will end with a little speculation on how the innovation could come about.

As we saw, a common use of progressive *know* follows the line of the structural-phenomenal distinction, where knowledge is construed not as something eternal, but as the impact of having that knowledge. For example, in the interview with the comic book creator, he notes "Meanwhile, I'm knowing they're probably going to bump me and go get Ridley Scott" (to direct). Having this knowledge causes Mark Steven Johnson not only to write the script, but make storyboards, drawings, music, etc. to bring a vision of how the movie should be along with the script. If he did this, then the producers would have no choice but to let him direct.

In other words, in the interview, Johnson is saying that what he knew at the time had an impact on his behaviour and strategy. This is the construal of pure knowledge into something more of the moment. Furthermore, the construal of *I know* to *that I know it* is no stretch whatsoever. It exists grammatically in the form of the verbal gerund. In more 'proper' English, he could have expressed

---

[26]But recall that drinking and eating predicates are rather tricky.

[27]In (Lambalgen & Hamm, 2005, p. 172-173) the transition from the accomplishment to the activity is an example of subtractive coercion, i.e. removing the goal-related information.

[28]This 'feeling' that no coercion is going on in the last example has implications for analysis. From (Croft & Cruse, 2003, p. 72): "The nature of our experience in many instances favors certain construals over others. It is these widespread typical or **default** construals that have led semanticists to posit a more rigid model of the mapping from linguistic meaning to experience." Note that then this is not purely a matter of convention but the cognitive ease in which certain construals can be made. However, it seems likely to be the case, that the 'easier' construals are much more amenable to be conventionalised.

[29]And even having its own OED entry.

[30]My personal aesthetic taste hopes that the causative use of *nervous* is confined to Big Brother viewers and vanishes along with the program.

himself along the lines, "Knowing they were going to hire Ridley Scott to direct, I not only wrote the script, but also mapped out the direction of the movie so that they had no choice but to hire me."

This construal of *know* already exists and is a common construction, but, I would imagine, that in the progressive construction the construal is much less readily accepted, and in standard English is certainly not acceptable. Nevertheless, I should hope that, by now, the reader's intuition about such examples have changed a bit, along with their notions about the status of interpretation. Mine certainly have.

<div align="right">

Appendix

</div>

This appendix is unfortunately not a full introduction to the technical aspects of the event calculus,[31] but rather, when combined with the Interlude, enough for the reader to get a feel and basic understanding of the formalisms used in this thesis. The first section shall list and also informally explain the axioms in detail. The second will demonstrate how these are not sufficient and shall sketch the computational system needed.

The Axioms

In the following, all variables are assumed to be universally quantified.[32]

Axiom 1. $Initially(f) \land \neg Clipped(0, f, t) \rightarrow HoldsAt(f, t)$

Axiom 2. $Happens(e, t) \land Initiates(e, f, t) \land t < t' \land \neg Clipped(t, f, t') \rightarrow HoldsAt(f, t')$

Axiom 3. $Happens(e, t) \land Initiates(e, f_1, t) \land t < t' \land t' = t + d \land Trajectory(f_1, t, f_2, d) \land \neg Clipped(t, f_1, t') \rightarrow HoldsAt(f_2, t')$

Axiom 4. $Happens(e, s) \land t < s < t' \land (Terminates(e, f, s) \lor Releases(e, f, s)) \rightarrow Clipped(t, f, t')$

Axiom 2 is really the embodiment of the principle of inertia. Suppose that a fluent f is initiated at time $t_1 > 0$, and that no 'f-relevant' event occurs between times t_1 and t_2.[33] Axiom 2 then says that f still holds at t_2. Note that since the predicate *Clipped* is defined both in terms of terminating and releasing events,

---

[31] The reader is recommended (Lambalgen & Hamm, 2005) for all of the gory details.

[32] Unless explicitly mentioned.

[33] 'f-relevant' can be seen formally in Axiom 4, which defines the predicate *Clipped*. An 'f-relevant' event is one that either terminates or releases a particular fluent f.

both the notions of instantaneous and continuous change are covered. For the former, the principle of inertia merely says that in the absences of relevant events no changes occur. But continuous change occurs due to force, not an instantaneous event, meaning that the absence of one of these events should not necessary imply the absence of change. The *Releases* predicate allows this to happen.[34]

Axiom 1 handles the analogous case where fluent holds initially. It implies that if a fluent holds at time 0 and no event has either terminated or released it before or at a time $t > 0$, it still holds at t. For Axiom 3, it is best to refer once again to our paradigm continuous change example. Let f_1 be instantiated by *filling*, and f_1 by *height(x)*. Should *filling* hold uninterrupted between from t until t', then for a certain x, *height(x)* with be true at t', with the value of x determined by the *Trajectory* predicate.

The completion

While the axioms listed and described above are a general account of inertia and change, they alone are not enough to enforce the concept of minimality discussed in the Interlude. This is best explained by using a simple scenario of turning a light on and off. For this, two events are needed – *switch-on* and *switch-off*, the fluents *light-on* and *light-off*, with the causal links being

\- $\neg HoldsAt(\text{light-on}, t) \rightarrow Initiates(\text{switch-on}, \text{light-on}, t)$

\- $\neg HoldsAt(\text{light-off}, t) \rightarrow Terminates(\text{switch-off}, \text{light-on}, t)$

Stated plainly, these mean that if the light is off, then turning the switch on initiates the light being on, and turning the switch off terminates the light being on (if the light is off). Now suppose we have a scenario that contains only the above causal information about light switches, plus the specific information

(5)   1. $Happens(\text{switch-on}, 5)$
        2. $Happens(\text{switch-off}, 10)$

In a minimal model, only two events occur – that of the light turned on at time 5 and off at time 10. Intuitively, then it is the case that the light is shining

---

[34]For example, in the bucket-filling example given in the Interlude, the event is that of turning the tap on. This event, by itself, does not change the state of the bucket, but instead begins the continuous force of water filling the bucket, as well as (via *Releases*) clips the parameterized bucket fluent, and thus allowing it to change. Without the releases predicate, if the height of the bucket was initially *height(0)* and the tap was turned on, the height of the water would remain 0 no matter how much water was poured into it. This is a circumstance that would only happen in an early Buñuel movie or if there was a hole in the bucket.

between times 5 and 10 and at no other times. Moreover, we have no information that anything holds initially. For example, suppose we want to check whether *?HoldsAt*(light-on,8). To do this we can apply Axiom 2, yielding

$Happens(switch\text{-}on, t) \land Initiates(switch\text{-}on, light\text{-}on, t) \land t < 8 \land \neg Clipped(t, light\text{-}on, 8)$
$\rightarrow$
$HoldsAt(light\text{-}on, 8)$

The scenario information says that the light is switched on at time 5, and since we know nothing about the light being on initially, we can assume that the event at time 5 does initiate the light being on. Moreover, no 'f-relevant' terminating events happen between time 5 and time 8 – we have to wait until time 10 for that to happen. By Axiom 2, then, the light is on at time 8,[35] and remains on until it is switched off at time 10.

But, there exist non-minimal models of the above scenario that can defy the above conclusion. For example, a non-minimal model of this scenario would add the events

- *Happens*(*switch-off*,7)

- *Happens*(*switch-on*,9)

In this extended model the light is no longer on at time 8, as the light has now been switched off at 7, and hence clipped between times 5 and 7. This is exactly what we don't want.[36] It could also be the case that a non-minimal model has information about *Initially*, e.g. *Initially*(*light-on*), which would mean that the light is on from time 0 and on continuously until turned off.[37]

Luckily, there is a method that will ensure that the only events listed in (5) are they only events that will occur in minimal models of this simple scenario. Moreover, the minimal model will also contain no *Initially* clauses as well.

This is done by a method called the *completion* of a predicate. Eliding the difficult to read formal definition, the idea is to circumscribe each predicate in the scenario (occurring as a the head of a clause)[38] that then allows for the most minimal conclusions that can be derived from the scenario information and the

---

[35]It is actually switched on at time 5.

[36]Instead, the scenario is expanded and one looks at minimal models of the expanded scenario.

[37]Note from the scenario that the *switch-on* event would have no effect at time five because *light-off* does not hold.

[38]The head of a clause is the consequent of a condition or an atomic scenario statement such as *Happens*(*Switch-on*, 5).

axioms.[39] Practically, this is done by defining every predicate[40] in terms of a biconditional that describes, via a series of disjunctions, every scenario relevant fact about the predicate. The syntactic definition is incredibly complex,[41] but can be easily illustrated with the *Happens* predicated in this scenario as they are only atomic statements. The general recipe is to start by gathering all clauses in which our predicate appears as a consequent. In this case it is just

- *Happens*(*switch-on*, 5)
- *Happens*(*switch-off*, 10)

We then turn the time parameter into a constraint, by turning the above into conditionals,[42]

$t = 5 \rightarrow Happens(switch\text{-}on, t)$
$t = 10 \rightarrow Happens(switch\text{-}off, t)$

The next step is to replace the terms in t_i[43] by a new set of variables and add the conjunct
$x = t_i \rightarrow p^i(x)$ to the body (antecedent) of each clause containing the predicate.[44] In this case the body is only the time constraint. So, these equalities are added to the antecedents of the 'variablised' above clauses, e.g. $x = switch\text{-}on \wedge t = 5 \rightarrow Happens(x, t)$.

Then (via further syntactic processes) a biconditional[45] is formed with one end being the a predicate and the other being the disjunctions of the bodies of the clauses (with added conjunct) in question. In this case, the process yields

$Happens(x, t) \leftrightarrow (x = switch\text{-}on \wedge t = 5) \vee (x = switch\text{-}off \wedge t = 10)$

which means that *Happens*(*switch-off*, 7) and *Happens*(*switch-off*, 9) do not happen, as the above formula says that the only two events that occur are the ones in (5).

But the analogous procedure needs to be done on the predicate *Initially*. This

---

[39]We did this above in a rather informal manner. In general, using the notion of reasoning backwards from initial query to sub-goals until we can't go any further, we can determine the times the derivable times that the various predicates are true.

[40]That exists as the head (atomic predicate of consequent) of a scenario clause.

[41]For those with a strong stomach, see (Lambalgen & Hamm, 2005, p.55).

[42]This is serendipity. The clauses could already have antecedents.

[43]These are terms in the language, such as *switch-on*.

[44]Each p is an instance of the predicate we are completing. Here, e.g. $p^1(x)$ is *Happens*(*switch-on*, t) and t_1 is *switch-on*. Also, technically, x is really $\overline{x}$ (same for t) which represents a possible string of variables that replace terms t in those predicates with arity greater than one (e.g. *Trajectory* or *Initiates*). Thankfully, this doesn't apply here.

[45]We will ignore the quantifiers here.

completion works at bit more simply as we have no information whatsoever about the predicate in the above scenario. The first step, then, here is to make a conditional with *falsum* as the antecedent:[46]

$\perp \rightarrow Initially(light - on)$

The biconditional for the completion is then

$\perp \leftrightarrow Initially(light - on)$

coupled with a clause that states $\neg HoldsAt(light-on, t) \rightarrow HoldsAt(light-off, t)$ it is ensured that the light is off initially.

The two completions, taken together, ensure that the temporal profile of the light's state is exactly the desired one sketched above.

---

[46]As the *Initially* predicate has no explicit time parameter, we only need to worry about the terms that could occur as arguments of the predicate in the scenario.
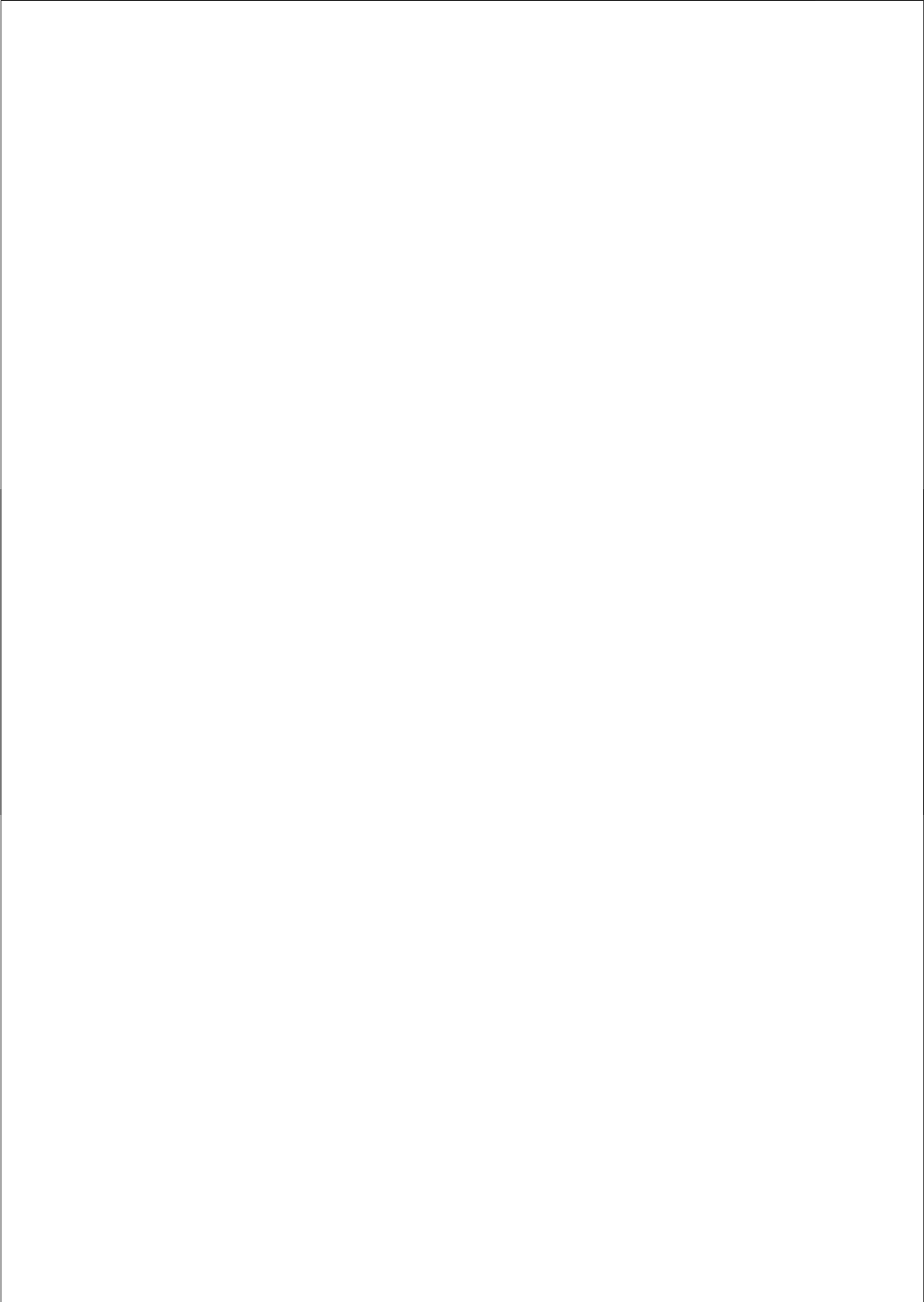
Bibliography

References

Abbot, V., Black, J., & Smith, E. (1985). The representation of scripts in memory. *Journal of Memory and Language*, *24*, 179-199.

Abney, S. (1987). *The English Noun Phrase in Its Sentential Aspect*. Unpublished doctoral dissertation, MIT.

Akmajian, A. (1977). The complement structure of perception verbs in an autonomous syntax framework. In *Formal syntax* (p. 427-460). Academic Press.

Anderson, S., & Conway, M. (1993). Investigating the structure of autobiographical memories. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *19*, 1178-1196.

Baker, C. (1995). *English syntax*. MIT Press.

Barsalou, L. (1988). the content and organization of autobiographical memories. In U. Neisser & E. Winograd (Eds.), *Remembering reconsidered: Ecological and traditional approaches to the study of memory*. Cambridge University Press.

Behrens, H. (2001). Cognitive-conceptual development and the acquisition of grammatical morphemes: the development of time concepts and verb tense. In M. Bowerman & S. Levinson (Eds.), *Language acquisition and conceptual development*. Cambridge University Press.

Bennett, M., & Partee, B. (1978). *Toward the logic of tense and aspect in English*. Bloomington Indiana Linguistics Club.

Binnick, R. (1991). *Time and the verb*. Oxford University Press.

Boland, A. (2006). *Aspect, tense and modality: Theory, typology, acquisition*. Unpublished doctoral dissertation, LOT, University of Utrecht.

Bower, G., Black, J., & Turner, T. (1979). Scripts in memory for text. *Cognitive Psychology*, *11*, 177-220.

Bresnan, J., & Nikitina, T. (2003). On the gradience of the dative alternation.

Bybee, J., Perkins, R., & Pagliuca, W. (1994). *The evolution of grammar: Tense, aspect, and modality in the languages of the world.* University of Chicago Press.

Callaway, M. (1901). The appositive participles in Anglo-Saxon. *Publications of the Modern Language Association of America, 16.*

Carey, S., & Xu, F. (2001). Infants' knowledge of objects: beyond object files and object tracking. *Cognition, 80,* 179-213.

Carlson, G. (1980). *Reference to kinds in English.* Garland.

Christensen, L. (2003). The acquisition of tense. In G. Josefsson (Ed.), *The acquisition of Swedish grammar* (p. 31-74). John Benjamins.

Cole, P. (1983). The grammatical role of the causee in universal grammar. *Internation Journal of American Linguistics, 49,* 115-133.

Comrie, B. (1976a). *Aspect.* Cambridge University Press.

Comrie, B. (1976b). Syntax and semantics 6: The grammar of causative constructions. In M. Shibatani (Ed.), (p. 261-312). Academic Press.

Comrie, B. (1985a). Language typology and syntactic description. In (p. 309-348). Cambridge University Press.

Comrie, B. (1985b). *Tense.* Cambridge University Press.

Comrie, B. (1989). *Language universals and linguistic typology (second edition).* University of Chicago Press.

Comrie, B., & Thompson, S. (1985). Lexical nominalization. In T. Shopen (Ed.), *Linguistic typology and syntactic description (vol. III)* (p. 349-398). Cambridge University Press.

Croft, W. (1991). *Syntactic categories and grammatical relations: The cognitive organization of information.* University of Chicago Press.

Croft, W. (1995). Modern syntactic typology. In M. Shibitani (Ed.), (p. 85-143). Oxford University Press.

Croft, W. (1998). The structure of events. In *The new psychology of language.* Lawrence Erlbaum.

Croft, W. (2000a). *Explaining Language Change.* Longman.

Croft, W. (2000b). *Verbs: aspect and argument structure.* Draft.

Croft, W. (2001). *Radical construction grammar.* Oxford University Press.

Croft, W., & Cruse, A. (2003). *Cognitive linguistics.* Cambridge University Press.

Cusack, B. (Ed.). (1988). *Everyday english 1500-1700: a reader.* Edinburgh University Press.

Dahl, O. (1981). On the definition of the telic-atelic (bounded-nonbounded) distinction. In *Tense and aspect. volume 14 of syntax and semantics.* Academic Press.

Davidson, D. (1967). The logical form of action sentences. In N. Rescher (Ed.), *The logic of decision and action* (chap. The Logical Form of Action Sentences). University of Pittsburgh Press.

Davies, J. (1981). *Kobon.* North-Holland Publishing Company.

Denison, D. (1983). *English Historical Syntax.* Longman.

Dik, S. (1997). *The theory of functional grammar. part 1: The structure of the clause.* Mouton de Gruyter.

Dixon, R. (2005). *A semantic approach to english grammar (second edition).* Oxford University Press.

Fanego, T. (2004). On reanalysis and actualization in syntactic change (the rise and development of English verbal gerunds). *Diachronica, 21*(1), 5-55.

Feest, S. van der, & Hout, A. van. (2002). Tense comprehension in child dutch. In *Bucld 26 proceedings* (p. 734-745).

Felser, C. (1999). *Verbal complement clauses : a minimalist study of direct perception constructions.* Benjamins.

Fillmore, C. (1963). The position of embedding transformations in a grammar. *Word, 19.*

Foley, W. (2005). Semantic parameters and the unnacusative split in the austronesian language family. *Studies in Language, 29*(2), 385-430.

Frawley, W. (1992). *Linguistic semantics.* Lawrence Erlbaum.

Freed, A. (1979). *The semantics of English aspectual complementation.* Reidel.

Garey, H. (1957). Verbal aspects in French. *Language, 33*, 91-110.

Gathercole, V. (1986). The acquisition of the present perfect : explaining difference in the speech of Scottish and American children. *Journal of Child Language, 13*, 537-560.

Gee, J. P. (1977). Comments on the paper by Akmajian. In *Formal syntax* (p. 461-482). Academic Press.

Gergely, G., & Csibra, G. (1993). Teleological reasoning in infancy: The naive theory of rational action. *Trends in Cognitive Sciences*(287-292).

Goldberg, A. (1995). *Constructions: A construction grammar approach to argument structure.* University of Chicago Press.

Goldberg, A. (2000). Patient arguments of causative verbs can be omitted: The role of information structure in argument distribution. *Language Science.*

Goldsmith, J., & Woisetschlaeger, E. (1982). The logic of the English progressive. *Linguistic Inquiry.*

Hamm, F., Kamp, H., & Lambalgen, M. (2006). There is no opposition between formal and cognitive semantics. *Theoretical Linguistics.*

Hamm, F., & Lambalgen, M. (2003). Event calculus, nominalization, and the progressive. *Linguistics and Philosophy, 26*(381-458).

Harris, A., & Campbell, L. (1995). *Historical syntax in cross-linguistic perspective.* Cambridge University Press.

Haspelmath, M. (1998). Does grammaticalization need reanalysis? *Studies in Language, 22*(315-351).

Heine, B., Ulrike, C., & Hunnemeyer, F. (1991). *Grammaticalization: A conceptual framwork.* University of Chicago Press.

Heyvaert, L. (2003). *A cognitive-functional approach to nominalization in English.* Mouton De Gruyter.

Higginbotham, J. (1983). The logic of perceptual reports: An extensional alternative to situation semantics. *Journal of Philosophy, LXXX*, 100-127.

Higginbotham, J. (2003). Remembering, imagining, and the first person. In A. Barber (Ed.), *Epistemology of langauge* (p. 496-533). Oxford University Press.

Higginbotham, J. (2005). Event positions: Suppression and emergence. *Theoretical Linguistics, 31*, 349-358.

Hindsill, D. (2001). *On the turning verbs into nouns.* Unpublished master's thesis, Universiteit van Amsterdam.

Hopper, P., & Traugott, E. (1993). *Grammaticalization.* Cambridge University Press.

Houston, A. (1989). The English gerund: Syntactic change and discourse function. In R. Fasold & D. Schriffin (Eds.), *Language change and variation* (p. 173-196). John Benjamins.

Hout, A. van. (1998). On the role of direct objects and particles in learning telicity in on the role of direct objects and particles in learning telicicity in Dutch and English. In A. Greenhill, M. Hughes, H. Littlefield, & H. Walsh (Eds.), *Proceedings of 22th boston university conference on language development.*

Huddleston, R. (2002a). Non-finite and verbless clauses. In G. Pullum & R. Huddleston (Eds.), *The Cambridge Grammar of the English language* (p. 1173-1268). Cambridge University Press.

Huddleston, R. (2002b). The verb. In *The Cambridge Grammar of the English language.* Cambridge University Press.

Irwin, B. (1967). *The development of the ING ending of the verbal noun and the present participle from c.700 – c.1400.* Unpublished doctoral dissertation, University of Wisconsin.

Jack, G. (1988). The origins of the English gerund. *NOWELE, 12*, 15-75.

Jackendoff, R. (1977). *X-bar syntax.* MIT Press.

Jackendoff, R. (1992). *Languages of the mind.* MIT Press.

Josephs, L. (1975). *Palau reference grammar.* University of Hawaii Press.

Kemmer, S., & Verhagen, A. (1994). The grammar of causatives and the conceptual structure of events. *Cognitive Linguistics, 5*(2), 115-156.

Kenny, A. (1963). *Action, emotion and will.* Routledge.

Klein, W. (1994). *Time in language.* Routledge.

Labov, W. (1989). The child as linguistic historian. *Language Variation and Change, 1*, 85-97.

Lambalgen, M. (2002). Logical constructions suggested by vision.

Lambalgen, M., Baggio, G., & Hagoort, P. (2007). Language, linguistics and cognition. In M. Stockhof & J. Groenendijk (Eds.), *Handbook of philosophical logic.* Springer.

Lambalgen, M., & Hamm, F. (2005). *The proper treatement of events.* Blackwell.

Langacker, R. (1977). Mechanisms of syntactic change. In (p. 57-139). University of Texas.

Lees, R. (1960). *The grammar of English nominalizations.* Mouton.

Levin, B. (2000). Aspect, lexical semantic representation, and argument expression. In *Proceedings of the 26th annual meeting of the Berkeley Linguistics Society* (p. 413-429).

Levin, B., & Rappaport-Hovav, M. (1995). *Unaccusativity: at the syntax-lexical semantics interface.* MIT Press.

Levin, B., & Rappaport-Hovav, M. (2004). The semantic determinants of argument expression: A view from the English resultative construction. In J. Gueron & J. Lecarme (Eds.), *The syntax of time.* MIT Press.

Lichtenstein, E., & Brewer, W. (1980). Memory for goal-directed events. *Cognitive Psychology, 12,* 412-445.

McCawley, J. (1976). Remarks on what can cause what. In M. Shibatani (Ed.), *Syntax and semantics 6: The grammar of causative constructions,* (p. 117-129). Academic Press.

McKoon, G., & MacFarland, T. (2001). Event templates in the represenataion of verbs. *Cognitive Psychology.*

Meltzoff, A. (1995). Understanding the intentions of others: Reenactment of intended acts by 18-month-old children. *Developmental Psychology, 31,* 838-850.

Miller, G., & Johnson-Laird, P. (1976). *Language and perception.* Harvard University Press.

Moens, M., & Steedman, M. (1988). Temporal ontology and temporal reference. *Computational Linguistics.*

Newtson, D. (1973). Attribution and the unit of peception of ongoing behavior. *Journal of Personality and Social Psychology, 28,* 28-38.

Newtson, D. (1993). The dynamics of action and interaction. In *A dynacmic systems approach to developmental: Applications.* MIT Press.

Olsen, M., & Weinberg, A. (1999). Innateness and the acquisition of grammatical aspect via lexical aspect. In *Proceedings of the Boston university conference on lanuguage development.*

Oxford english dictionary. (1989). Oxford University Press.

Parsons, T. (1990). *Events in the semantics of English.* MIT University Press.

Pawley, A. (1987). Encoding events in kalam and english: different logics for reporting experience. In R. Tomlin (Ed.), *Coherence and grounding in discourse.* John Benjamins.

Poutsma, H. (1923). *The infinitive, the gerund and the participles of the English verb.* P. Noordhoff.

Pullum, G. (1991). English nominal gerund phrases as noun phrases with verb-phrase heads. *Linguistics, 29,* 763-399.

Pullum, G., & Scholz, B. (2002). Empirical assessment of stimulus poverty arguments. *The Linguistic Review, 19.*

Quine, W. (1996). Events and reification. In R. Casati & A. Varzi (Eds.), *Events.* Aldershot.

Quirk, R., Greenbaum, L., & Svarvik. (1985). *A comprehensive grammar of the English language.* Longman Linguistics Library.

Rappaport-Hovav, M., & Levin, B. (2001). An event structure account of english resultatives. *Language, 77*(4).

Reuland, E. (1983). Governing-ing. *Linguistic Inquiry, 14*, 10-34.

Rosch, E. (1978). Principles of categorization. In E. Rosch & B. LLoyd (Eds.), *Cognition and categorization.* Erlbaum.

Rosch, E., Mervis, C., Gray, W., Johnson, D., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology, 8*, 382-439.

Rothstein, S. (2004). *Structuring events: a study in the semantics of lexical aspect.* Blackwell.

Ryle, G. (1949). *The concept of mind.* Hutchinson.

Schachter, P. (1976). A nontransformational account gerundive nominals in English. *Linguistic Inquiry, 7*, 205-241.

Shipley, E., & Shepperson, B. (1990). Countable entities: developmental changes. *Cognition, 34*, 109-136.

Slobin, D. (1985). Crosslinguistic evidence for the language-making capacity. In *The crosslinguistic study of language acquisition, vol. 2: Theoretical issues.* Erlbaum.

Slobin, D. (2001). Form-function relations: how do children find out what they are? In M. Bowerman & S. Levinson (Eds.), *Language acquisition and conceptual development.* Cambridge University Press.

Song, J. (2001). *Linguistic typology: morphology and syntax.* Longman Linguistics Library.

Steedman, M. (n.d.). *The productions of time.*

Tajima, M. (1985). *The syntactic development of the gerund in Middle English.* Nan'un-do.

Timberlake, A. (1977). Mechanisms of syntactic change. In (p. 141-177). University of Texas.

Tomasello, M. (1999). *The cultural origins of human cognition.* Harvard University Press.

Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition.* Harvard University Press.

Tomasello, M., & Brooks, P. (1999). How children constrain their argument structure constructions how children constrain their argument structure constructions how children constrain their argument structure constructions. *Language, 75*(4), 720-738.

Trabasso, T., & Stein, N. (1994). Using goal-plan knowledge to merge the past with the present and the future in narrating events. In *The development of future-oriented processes.* University of Chicago Press.

Traugott, E., & Dasher, R. (2002). *Regularity in semantic change.* Cambridge University Press.

Travis, L. (1997). Goal-based organization of event memory in toddlers. In

P. Bauer & T. Bovig (Eds.), *Developmental spans in event comprehension and representation: Bridging fictional and actual events.* Erlbaum.

Valin, R. J., & La Polla, R. (1997). *Syntax : structure, meaning and function.* Cambridge University Press.

Vendler, Z. (1967). *Linguistics in philosophy.* Cornell University Press.

Vendler, Z. (1968). *Adjectives and nominalizations.* Mouton.

Visser, F. (1963-1973). *An historical syntax of the English lanugage.* E.J. Brill.

Wagner, L. (2001). Aspectual influences on early tense comprehension. *Journal of Child Language, 28*, 661-681.

Wagner, L. (2002). Understanding completion entailments in the absence of agency cues. *Journal of Child Language, 29*, 109-125.

Wagner, L., & Carey, S. (2003). Individuation of objects and events: a developmental study. *Cognition, 90*, 163-191.

Weist, R. (1991). Spatial and temporal location in child language. *First Language, II*, 253-267.

Weist, R., Lyytinen, P., Wysocka, J., & Atanassova, M. (1997). The interaction of language and thought in children's langauge acquisition. *Journal of Child Language, 24*, 81-121.

Weist, R., Wysocka, H., & Lyytinen, P. (1991). A cross-linguistic perspective on the development of temporal systems. *Journal of Child Language, 18*, 67-92.

Wittgenstein, L. (1953). *Philosophical investigations.* Oxford: Blackwell.

Zacks, J., & Tversky, B. (2001). Event structure in perception and conception. *Psychological Bulletin, 127*(1), 3-21.

Zacks, J., Tversky, B., & Iyer, G. (2001). Perceiving, remembering, and communicating structure in events. *Journal of Experimental Psychology: General, 130*(1), 29-58.
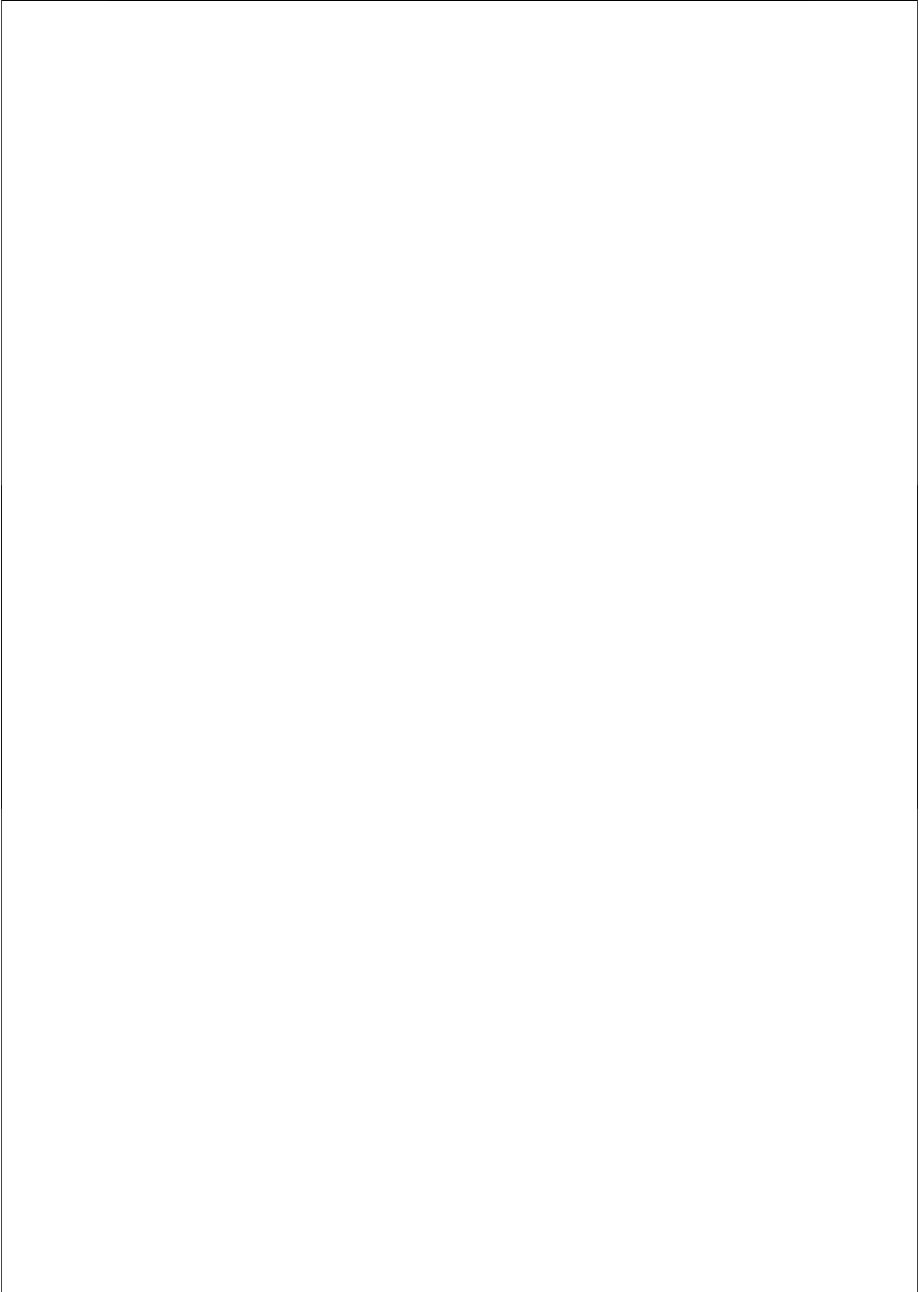
Samenvatting

Dit proefschrift bevat een onderzoeking van enkele onderwerpen over de semantiek van gebeurtenissen met als doel een uitbreiding van het formele systeem genoemd *Event Calculus* (Lambalgen & Hamm, 2005), een computationele behandeling van tijd en aspect. Ik bespreek werkwoorden van perceptie en causatieve constructies en bepaal hun temporele en implicationele eigenschappen.

Ik geef een nieuwe behandeling van de mogelijke *Aktionsarten* voor *see* and *watch* en de vele bestaande causatieve configuraties. Zij worden geformaliseerd en breiden de scope van de *Event Calculus* uit met de structuren van coördinatie van veelvoudige gebeurtenissen. Het Engels *Ing* nominalisatie wordt ook grondig besproken en ik geef een formalisatie van de vele interpretaties van verbale gerundia als processen.

Het bovenste gaat gepaard met een empirisch én theoretisch onderzoek van de verschuiving van perspectief in uitingen die de bovengenoemde onderwerpen bevatten. Ik laat zien dat zinnen die in de literatuur als niet grammaticale worden bestempeld eigenlijk vaak worden gebruikt door eerste taal sprekers van het Engels. De door mij ontwikkelde formele scenario's met de gegeven computationele gereedschap van de *Event Calculus* tonen aan hoe de 'nieuwe' betekenissen (dat is, de nieuwe *Aktionsart*, aspectuele eigenschappen en logische implicaties) voorspeld kunnen worden.

Hoofdstuk 1 geeft een algemene inleiding, leesgids en plan van de dissertatie.

Hoofdstuk 2 is een overzicht van *Aktionsart*, hiërarchische structuur van gebeurtenissen en de verschillenden soorten van coercie.

Hoofdstuk 3 geeft een overzicht van de relevante psycho-linguistieke literatuur verbonden met de discussie in hoofdstuk 2. De eerste helft bespreekt de cognitieve grondslagen van de structuur van gebeurtenissen en hiërarchische organisatie die de conceptuele typologie van hoofdstuk 2 weerspiegelt. De tweede helft onder-

307

zoekt vele resultaten die de acquisitie van tijd en aspect van kinderen betreffen. Dit culmineert in een verband tot het werven van de structuur van gebeurtenissen.

Deze twee gedeelte fungeren/dienen als psychologische grondslag van de formalisatie van de *Event Calculus*. Het eerste gedeelte geeft argumenten voor de notie van *default scenarios* en het tweede gedeelte voor de computationele gereedschap die de interpretatie van tijd en aspect berekent.

Hoofdstuk 4 onderzoekt de syntaxis, semantiek en logische eigenschappen van de twee klassen van perceptie-werkwoorden, namelijk die van *see* en *watch*. Voor beide klassen worden verschillende soorten van *Aktionsarten* en temporele profielen bepaald. Ik stel dan enkele 'templates' voor om de tijd en aspectuele interacties te formaliseren. De in de template gedefinieerde semantische 'constraints' laten zien hoe vermoedelijke ongrammaticale perceptie zinnen een juiste interpretatie kunnen krijgen.

Hoofdstuk 5 is een studie van causatieve constructies. Het begint met een crosstallige onderzoek om de relevante semantische eigenschappen te bepalen (inclusief verschillende causale niveaus en rollen van de participanten). Met deze inzichten worden vele engelse causatieve constructies bekeken; daaruit komen verrassende resultaten wat betreft de causatieve 'alternation' en de perifrastische causatieven. Deze resultaten worden dan gebruikt voor het formaliseren van de causatieve templates. De templates voorspellen én de standaard interpretatie én de voorbeelden veroorzaakt door de verschuiving van perspectief.

Hoofdstuk 6 onderzoekt de engelse *Ing* nominalisatie vanuit een synchronisch en een diachronisch perspectief. Ik bespreek eerst de voorafgaande analysen en dan de geschiedenis van deze vorm. Dit leidt tot vragen over haar synchronische analyse (syntactisch en semantisch). Ik presenteer empirische voorbeelden van vermoedelijke syntactische onmogelijkheden en semantische anomalieën. Ik concludeer dan dat de oorspronkelijke formalisatie van de *Event Calculus* makkelijk uitgebreid kan worden door eenvoudige modificaties van de scenario's. Uiteindelijk formaliseer ik verschillende processen en gebeurtenis-achtige lezingen van verbale nominalen.

Hoofdstuk 7 bevat tenslotte een samenvatting van de belangrijke resultaten en conclusies.