TENSE AND ASPECT

OLIVER BOTT & FRITZ HAMM

Time flies like an arrow. Fruit flies like a banana.
Groucho Marx

ABSTRACT. Many languages have grammatical means to signal the time when an action, or an event occurs, a state or process holds. This phenomenon is called tense. For example in English the morpheme -ed attached to a verb like walk forming walked conveys the information that the event denoted by the verb occurred before the present time.

The notion aspect refers to the internal constituency of actions, events, states, processes or a situation. For instance it may indicate that an action is completed or still ongoing. English typically uses the -ing form of verbs to indicate ongoing processes as in He is building a house.

After a short introduction of basic notions concerning tense and aspect we will briefly address temporal logic, then we will explain Reichbach’s famous distinction between speech time, event time and reference time. The section will be followed by the most important observations concerning lexical and grammatical aspect. This part also contains the famous imperfective paradox. Next we introduce two widely discussed theories of temporality and show how these theories cope with the imperfective paradox. Event semantic theories treat events as ontological primitives. In the following section we therefore show how time can be constructed from event structures as exemplified by the Russell-Kamp construction. This section is followed by a section that shows how temporal information is expressed in Artificial Intelligence. We use the event calculus from AI to present a solution of the imperfective paradox as an instance of the frame problem. The last section is devoted to gathering psycholinguistic evidence showing that at least some of the philosophical and semantic concepts discussed so far may be cognitively real.

CONTENTS

1. Introduction 2
2. Temporal Logic 4
4. Lexical and Grammatical Aspect 7
5. Intensionality versus Events 9
6. From Events to Time 12
7. Event Calculus 14
8. Cognitive aspects 21
8.1. Lexical Aspect 22
8.2. Grammatical Aspect 23

Date: March 13, 2013.
1. INTRODUCTION

_Tense_ roughly means reference to the time events take place, processes or states hold. English, for instance clearly distinguishes between past and non-past tense as in (1-a) and (1-b) and (1-c).

(1) a. John promised to pay ten pounds.
   b. I promise to pay you ten pounds.
   c. The student will work on his thesis.

In English typically _verb_ forms signal the time when an action or events occurs or a state holds. Thus in (1-a) John’s promise was given before the present time, in (1-b) the promising is simultaneous with the present time and (1-c) says that the student’s work on his thesis will occur at some time after the present. It should be noted that the verb forms used to express temporal information may also be used to signal information that is not purely temporal. For instance the present tense form in _John walks_ characterizes a certain habit of John. And in the statement _The train departs at five o’clock tomorrow_ the present verb form clearly has a futurate meaning. This is the reason why English distinguishes between past and non-past but not between future and non-future. Moreover, many linguists exclude future as a pure tense, because the auxiliary _will_ may be used to express volition as in _He will go swimming in dangerous water_.

Apart from _absolute_ tense – exemplified by (1-a) – (1-c) – where the reference point from which the location in time is evaluated is the present there is also _relative_ tense where the reference point from which the location in time is evaluated is not necessarily the present but may be given by context. Thus in a sentence like _The student had worked on his thesis_ the student’s work took place at a time before a reference point in the past in contrast to (1-c) where the reference point is the present. Other instances of relative tense in English are the present perfect and the future perfect in (2-a) respectively (2-b).

(2) a. I have lost my watch.
   b. The student will have finished his thesis next month.

The notion _aspect_ according to Comrie refers to “[the] different ways of viewing the internal temporal constituency of a situation” Comrie [1976], p. 3. It is customary to distinguish between _lexical_ and _grammatical_ aspect. The following are examples concerning lexical aspect. Since Vendler [1967] linguists distinguish at least four aspectual classes. These are states like _know_, activities like _run_, accomplishments like _cross the street_ and achievements exemplified by a verb like _recognize_. Accomplishments describe the internal temporal constituency of a situation in a rather

---

1For a thorough discussion of these issues we refer the reader to Comrie [1985].
more detailed way than say activities. Crossing the street for example includes the starting of a crossing activity which goes on for some time and involves a result state that is characterized by being on the other state of the street. No such elaborate descriptions are necessary for activities or achievements. There are also grammatical means to distinguish between these aspectual classes; for instance states and achievements cannot occur with the progressive while activities and accomplishments can; more about this below in section 4. These differences are customarily considered as lexical differences and therefore this area is dubbed lexical aspect.

A further crucial aspectual difference is that between perfective and imperfective aspect. Comrie characterizes these notions as follows.

...perfectivity indicates the view of the situation as a single whole, without distinction of the various separate phases that make up that situation, while the imperfective pays essential attention to the internal structure of the situation.

Comrie [1976], p. 16

In English this difference is often expressed by grammatical means, for instance by past tense versus past progressive. This is an instance of grammatical aspect. Slavic language often have an elaborate grammatical system to signal the difference between imperfectivity and perfectivity. Here is an example from Russian.

(3) a. Petr pisal,/*napisal pis’mo, kogda on byl prervan telefonym zvonkom.
   Peter was-writing/*wrote letter, when he was interrupted by-the telephone call.

b. Petr *pisal,/*napisal pis’mo i otnes ego na poˇctu.
   Peter *was-writing/wrote letter and brought it to post-office.

Sentence (3-b) presents the writing of a letter as a single whole and it is clear that this activity was finished it was brought to the post office. By contrast (3-a) focuses on an ongoing activity of letter writing in the past which is not completed because of the interruption. In fact it is consistent with the truth of (3-a) that the letter was never completely written.

The natural language categories tense and aspect embody the linguistic encoding of time. From a typological point of view these categories are typical verbal categories. This means that if these categories are morphologically realized in a language then these morphemes attach to the verb. Verbs usually express events, processes, actions or states and the temporal morphems locate these eventualities in time. Of course this does not mean that languages that lack such morphemes are not able to express temporal relations. They just have to choose other means for this purpose. One famous example of a language not containing temporal morphology is Mandarin Chinese. Often tense and aspect cannot be clearly separated.

2But see Lecarme [2008] for possible exceptions.

3But note that verbs, verb phrases or sentences may be nominalized and this nominalization process may transfer verbal properties to nominals. We will not discuss these linguistic subtleties here.
For example the Quiché prefix x- denotes a completed (aspect) action in the past (tense). Typologists also observed an asymmetry between past and future. Most languages that possess temporal morphology use these means to differentiate between past and non-past. There are few languages which use temporal morphology to distinguish between future and non-future.

The body of this paper is, however, not devoted to linguistic subtleties concerning tense and aspect but to a brief and necessarily incomplete presentation of major philosophical and semantic theories of these notions.

2. **Temporal Logic**

Temporal logic was introduced by Arthur Prior (see for instance Prior [1967]). Here we will sketch only propositional temporal logic (for a more thorough introduction the reader is advised to consult Gamut [1991], chapters 2 and 3) and the SEP entry on temporal logic. Similar to the modal operators □ and ◊ temporal logic introduces operators G, H, F and P. We summarize the intuitive meaning of these operators in Table 1.

<table>
<thead>
<tr>
<th>Operator</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gφ</td>
<td>it is always going to be the case that φ</td>
</tr>
<tr>
<td>Hφ</td>
<td>it always has been the case that φ</td>
</tr>
<tr>
<td>Fφ</td>
<td>it will at some stage in the future be the case that φ</td>
</tr>
<tr>
<td>Pφ</td>
<td>it was at some stage in the past that φ</td>
</tr>
</tbody>
</table>

By adding the above operators to propositional logic we extend propositional logic to propositional temporal logic. Let q abbreviate the sentence *Sam is working*, then we are able to express the following verb tenses in temporal logic:

(4) a. q  Sam is working.
    b. Fq  Sam will work.
    c. Pq  Sam worked.
    d. PPq  Sam had worked.
    e. FPq  Sam will have worked.
    f. PFq  Sam would work.

A model M for propositional temporal logic consists of a nonempty set T of moments of time, an earlier than relation ≤ and a valuation function I which for each moment of time and each proposition letter q assigns a truth value I_M,t(q). The following definition characterizes I_M,t(φ) for the additional temporal operators only.

**Definition 1.** Let M be a model, then I_M,t(φ) is defined by the following clauses:

1. I_M,t(Gφ) = 1 if for all t’ ∈ T such that t ≤ t’ : I_M,t’(φ) = 1.
2. I_M,t(Hφ) = 1 if for all t’ ∈ T such that t’ < t : I_M,t’(φ) = 1.
3. I_M,t(Fφ) = 1 if there exists t’ ∈ T such that t < t’ : I_M,t’(φ) = 1.
4. I_M,t(Pφ) = 1 if there exists t’ ∈ T such that t’ < t : I_M,t’(φ) = 1.

---

^4For further evidence from Hausa and Chichewâ see Spencer [1991], chapter 1.3.
These clauses are analogous to clauses which define the modal operators $\Box$ and $\Diamond - G$ and $H$ corresponding to $\Box -$ but instead of possible worlds the valuation function takes moments of time as arguments. Therefore principle (5) is valid in any model since the modal analogue is valid for $\Box$ independently of the accessibility relation.

(5) \[ G(\phi \to \psi) \to (G\phi \to G\psi) \] (similar for $H$).

However, principle (6) which is often considered as valid for $\Box$ is presumably not valid for $G$.

(6) \[ G\phi \to \phi \]

Formula (6) is equivalent to $\phi \to F\phi$. The latter formula says that if $\phi$ is the case then $\phi$ will be the case which is intuitively not correct. If we assume that $\prec$ is irreflexive, Principle (6) becomes invalid. This assumption makes sense since it means that no time point is earlier than itself. For extensions of temporal logic and more examples concerning the correspondence between properties of models and the validity or invalidity of formulas see Gamut [1991], chapter 2.

This concludes our brief remarks about temporal logic. We now turn to an alternative approach due to the philosopher Hans Reichenbach.

3. SPEECH TIME, EVENT TIME AND REFERENCE TIME

The notions speech time, event time, and reference time were introduced by Reichenbach [1947] in order to distinguish simple past and present perfect or more generally absolute and relative tense. According to Comrie [1985] who refined Reichenbach’s system in chapter 6 of his book on tense, speech time and event time are sufficient for the analysis of absolute time; i.e., present, (simple) past and future. But for relative tense - of which the present perfect is but one example - reference time is required. Let us explain these notions by applying them to distinguish between past and present perfect.

Consider the following examples from Steedman [1997].

(7) a. I have lost my watch.
    b. Yesterday, I lost my watch.

It has often been observed that the present perfect in English has present relevance. For instance the continuation of (7-a) with but I have found it again is infelicitous in English; the German translation of this sequence is acceptable, showing that the German perfect is more like a past tense, but see Kamp et al. [2011] for an extensive discussion. The same continuation is fine for sentence (7-b). In this sense the perfect is a relative tense and the past an absolute tense in English. Let $E$ be short for event time and let $S, R$ stand for speech time and reference time.
For the simple past both event time and reference time are situated on the time line before speech time. In case of the perfect R and S are simultaneous and E is earlier than both R and S. Intuitively reference time represents the perspective from which an eventuality is perceived on the time line. This is not a purely truth functional theory of tense like temporal logic but a pragmatic theory of tense. For instance we may assume that reference time in contrast to event time has to be mutually known by communication partners. By using the perfect reference time is known since it coincides with speech time. However, reference time is not necessarily known in case of the past since it is just required to be earlier than speech time and to coincide with event time. This may explain why sentences in the past sound strange when uttered out of the blue. Steedman [1997] presents the following examples:

(8) a. *Chapman breathed a sign of relief.
   b. When Nixon was elected, Chapman breathed a sigh of relief.

Sentence (8-a) in the past uttered out of the blue is infelicitous while sentence (8-b) is fine, since in this case the when-clause introduces the reference time.

This is be far not the complete story of the perfect. For example Comrie [1976] distinguishes four typical uses of the perfect, the perfect of result, the experiential perfect, the perfect of persistent situation and the perfect of recent past. For an extensive recent discussion the reader is advised to consult Kamp et al. [2011].

The following table summerizes the positions of event time, reference time and speech time for other tenses:

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>E, R, S</td>
<td>Sam is working.</td>
<td></td>
</tr>
<tr>
<td>S — E, R</td>
<td>Sam will work.</td>
<td></td>
</tr>
<tr>
<td>E, R — S</td>
<td>Sam worked.</td>
<td></td>
</tr>
<tr>
<td>E — R — S</td>
<td>Sam had worked.</td>
<td></td>
</tr>
<tr>
<td>S — E — R</td>
<td>Sam will have worked.</td>
<td></td>
</tr>
<tr>
<td>R — E — S</td>
<td>Sam would work.</td>
<td></td>
</tr>
</tbody>
</table>

Comrie [1985] and Gamut [1991] point out that this cannot be a complete account for temporal constructions that occur in natural languages. For instance the temporal profile of the sentence Sam would have worked cannot be analyzed by using a single reference time. This is one reason for the extended requirements for a formal theory of tense proposed by [Comrie, 1985, chapter six].

The next section will show that for an adequate analysis of tense and aspect the semantics of verbs, verb phrases and whole sentences has to be taken into account.
Linguists distinguish four or five aspectual classes or Aktionsarten. The following four were introduced by the philosopher Zeno Vendler in chapter four of his book Vendler [1967] on the basis of earlier work by philosophers such as Kenny, Ryle and Aristotle. For a formal definition of Aktionsart the reader is referred to van Lambalgen and Hamm [2005], p 85 ff.

Table 3

<table>
<thead>
<tr>
<th>States</th>
<th>Activities</th>
<th>Accomplishments</th>
<th>Achievements</th>
</tr>
</thead>
<tbody>
<tr>
<td>know</td>
<td>run</td>
<td>cross the street</td>
<td>recognize</td>
</tr>
<tr>
<td>be beautiful</td>
<td>swim</td>
<td>draw a circle</td>
<td>reach</td>
</tr>
<tr>
<td>believe</td>
<td>walk</td>
<td>paint a picture</td>
<td>find</td>
</tr>
</tbody>
</table>

In addition linguists often assume that verbs like flash, spot and blink form an extra class the class of semelfactives or points (see Smith [1991]).

A useful notion for distinguishing Aktionarten is the event nucleus introduced by Moens and Steedman [1988]. The event nucleus is constructed from a preparatory phase, a culminating event and a consequent phase. Activities only refer to the preparatory phase, states only to the consequent phase, achievements to the culminating event and to the consequent phase and finally accomplishments to all three parts of the event nucleus.

<table>
<thead>
<tr>
<th>preparation</th>
<th>event</th>
<th>consequent</th>
</tr>
</thead>
</table>

A formalisation of Aktionsarten based on the event nucleus is contained in van Lambalgen and Hamm [2005], p 85 ff. Semelfactives are defined here too. They are assumed to refer to the culminating event only.

Vendler also proposed several linguistic tests which allow to distinguish these verb classes. Here we will present only a small selection; for a much more comprehensive list the reader is referred to Dowty [1979]. The first test separates non-statives from statives. Only non-statives occur in the progressive.

(9) a. *Sam is knowing the answer.
    b. Sam is running.
    c. Sam is writing a letter.

Accomplishment verbs prefer for-adverbials as temporal modifiers whereas activity verbs allow only in-adverbials.

(10) a. ?Sam wrote a letter for an hour.
    b. Sam wrote a letter in an hour.
    c. Sam swam for an hour.
    d. *Sam swam in an hour.

Achievement verbs are usually infelicitous with for-adverbials but allow the combination with in-adverbials.

(11) a. Sam reached the summit in five hours.
b. *Sam reached the summit for five hours.

The last test we mention here concerns the different entailment patterns of activities and accomplishments in the progressive. Activities in the past progressive entail their past reading but accomplishments in the past progressive don’t.

(12)  a. Sam was pushing a cart. ⇒
      b. Sam pushed a cart.

(13)  a. Sam was building a house. ∅
      b. Sam built a house.

This is an instance of the famous imperfective paradox. Of course taken literally this is not a paradox. Nevertheless the entailment patterns (12) and (13) pose a significant problem for any formal theory of tense and aspect. For instance, the first formal proposal for truth conditions of sentences in the past progressive stated that a sentence S in the past progressive is true if and only if there exists an open interval before the speech time at which the sentence without the progressive is true. According to definition 1.4 this immediately validates pattern (12) but without qualifications this truth definition does not account for pattern (13). Sections 5 and 7 contain extensive discussion of the imperfective paradox.

Vendler thought of the Aktionarten as lexical properties of verbs. That this position is doubtful was pointed out by many linguists, in particular by Dowty [1979] and Verkuyl [1993]. For a simple case consider the verb drink. If we assume that being an activity is a lexical property of this verb then surely by combining it with the noun wine this property is preserved for the complex phrase drink wine. However, if we combine the verb drink with the noun phrase a bottle of wine we turn an activity into an accomplishment. The converse problem arises if we assume that being an accomplishment is a lexical property of drink. This phenomenon of aspectual reinterpretation was dubbed coercion in Moens and Steedman [1988].

Aktionsart is not even definitely fixed at the VP-level. The verb arrive in

(14)  a. Chapman arrived.
      b. *Chapman arrived all night.

is an achievement as demonstrated by the ungrammaticality of (14-b). But if we choose a bare plural as subject it is turned into an activity and sentence (15) is grammatical.

(15) Visitors arrived all night.

Therefore a final decision about aspectual class cannot be reached below the sentence level. Steedman’s famous sentence (16)

(16) It took me two years to play the “minute Waltz” in less than sixty seconds for one hour without stopping.
shows that coericon can be iterated and that aspectual class can switch back and forth in this iteration process. We close this section with two further examples of aspectual coercion. In the first one a stative verb is transformed into an activity.

The verb *resemble* is a stative verb and therefore (17-a) is unacceptable but (17-b) is fine.

(17) a. *She is resembling her mother.*
   b. She resembles her mother.

However, if we add *more and more every day* to (17-a) thus forming

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

we get an acceptable result. This is due to the phrase *more and more every day* which coerces a state into an activity. Our last example is about temporal modification with *for*-adverbials. As pointed out above accomplishments are usually rather bad with *for*-adverbials. They require *in*-adverbials. But consider sentence (19) from van Lambalgen and Hamm [2005].

(19) Pollini played Opus 111 for two weeks.

This sentence seems to be fine. However it is not interpreted as an accomplishment any more but is reinterpreted as an iterated activity. To get this reading a lot of non-linguistic knowledge is required. First one has to guess that Opus 111 refers to Beethoven’s last piano sonata and moreover one has to be aware that this piece takes about 25 minutes. Under this assumptions sentence (19) says that Pollini played Opus 111 repeatedly within a time span of two weeks. Formal accounts of coercion phenomena are contained in Egg [2005], van Lambalgen and Hamm [2005] and Steedman [1997] among others. The reader is also advised to consult Steedmans updated version of Steedman [1997], which is available from his homepage: http://homepages.inf.ed.ac.uk/steedman/. For processing studies of coercion are reviewed in Bott [2010]; see also Section 8.

5. INTENSIONALITY VERSUS EVENTS

The most influential approach for solving the imperfective paradox is due to Dowty [1979]. The basic idea is to treat the progressive as a modal operator (□) but restrict its domain to inertia worlds. The notion *inertia world* is characterized informally. An inertia world is exactly like the real world up to the time of evaluation, after this time point an inertia world may differ from the real world but nevertheless is assumed to be as similar as possible to the real world. Given this notion the following definition introduces truth conditions for sentences in the progressive.

**Definition 2.** \( \text{PROG}(\phi) \) is true with respect to interval \( I \) and world \( w \) if and only if \( \forall w' \in IE(w) \) there exists an interval \( I' \supseteq I \) such that \( \phi \) is true with respect to \( I' \) and \( w' \). In addition \( I \) should not be the final subinterval of \( I' \). \( IE(w) \) is the set of inertia worlds for \( w \).
Definition 2 does not yet account for the different entailment patterns of activities and accomplishments. Two additional assumptions are required. The first one concerns activities the second one accomplishments.

(20) If $\phi$ is an activity that is true with respect to interval $I$, then $\phi$ is true with respect to any sufficiently large subinterval of $I$.

Accomplishments like Mary draw a circle are split up into two parts, an activity part Mary draw that satisfies principle (20) and a result part which is characterized by the existence of a circle. Both parts are connected by a causality relation $\text{CAUSE}$.

(21) $\text{CAUSE}[\text{Mary draw}, \ A \text{ circle comes into existence}]$

Thus an accomplishment has a richer internal structure than an activity.

With these additional requirements the inference patterns for activities and accomplishments follow. Let us show that Mary was pushing a cart implies Mary pushed a cart. Assume that the first sentence is true in $w$ with respect to speech time $S$. Then Mary be pushing a cart is true in $w$ with respect to an interval $I$ before $S$. According to definition 2 Mary push a cart is then true in every inertia world $w' \in IE(w)$ with respect to an interval $I' \supseteq I$. Because of postulate (20) Mary push a cart ist true in every $w' \in IE(w)$ with respect to interval $I$ ($I$ is a subinterval of $I'$). Now the definition of inertia worlds implies that Mary push a cart is true in $w$ with respect to interval $I$. This means that Mary pushed a cart is true in $w$ with respect to speech time $S$.

A completely analogous argument shows that Mary was drawing a circle implies Mary drew. But postulate (20) only holds for the first part of accomplishments, the result part is only required to hold in every inertia world $w'$ with respect to a larger interval $I'$, normally one after speech time. Since the inertia worlds after $S$ may differ from $w$ and since the subinterval property (20) is not required to hold for the result parts of accomplishments Mary drew a circle is not implied.

Many researchers assume that inertia worlds introduce a notion of normality for the semantic analysis of the progressive. The thunderbolt in example (22) from Landman [1992] seems to break the normal development of the real world. Dowty’s analysis seems to be tailored for cases like this one.

(22) Mary was crossing the street, when a thunderbolt from heaven struck her down.

However, this normality interpretation of inertia leads to objections. The first one is attributed to Frank Flach in Ogihara [1989]. Consider sentence (23)

(23) John was crossing the street, when he was hit by a truck.

and a situation in which the truck is only a few centimeters away from John. Moreover the truck’s speed is such that it is impossible to stop before reaching John. In this situation the normal course of events is such that John will never reach the other side of the street. Therefore given definition 2 sentence (23) should be false under
these circumstances. This is intuitively incorrect. A similar objection was raised by Bonomi [1997].

However, a normality interpretation seems to be at least questionable. Consider the following example from Naumann and Piñón [1997].

(24) Rebecca was running across the minefield.

This does presumably not mean that when you are running across a minefield, you normally get to the other side eventually.

Note that all these objections use a particular informal interpretation of the notion *inertia*. In order to judge the validity of these objections to Dowty’s theory in a precise way therefore requires an explicit theory of this notion with exact empirical predictions.

Dowty analyzes the progressive as an intensional construction. Does the progressive satisfy the standard philosophical tests for intensional construction; i.e the invalidity of substitution of co-designative proper names and the impossibility to pull the existential quantifier out of the context created by the intensional construction (see Bealer and Mönnich [1989] for a more detailed discussion). Let us concentrate on accomplishments only and let us consider sentence (25).

(25) Jackson Pollock was painting Ruth Rendel.

Given that Ruth Rendel and Barbara Vine are codesignative proper names sentence (25) implies Jackson Pollock was painting Barbara Vine. Therefore the first test fails. Co-designative proper names are substitutable salva veritate in progressive constructions. The second one, however, applies, since

(26) Carlos was building a house.

does not imply that there is a house that Carlos was building. These tests therefore don’t provide a clear answer to the question whether the progressive is an intensional construction. Even if the answer is yes, the progressive certainly is a different type of intensional construction as, for instance, propositional attitudes like *believe* or *doubt*.

An extensional alternative was developed by Parsons in Parsons [1989] and Parsons [1990]. Parsons like Davidson [1967] assumes that first-order quantification over events is possible. But unlike Davidson Parsons assumes quantification over eventualities in the sense of Bach [1986] which include events proper but also states, processes etc.. Like Dowty Parsons supposes that eventualities are split up in a development phase and a culmination phase. In order to represent this difference in first-order logic two new predicates are introduced, $\text{Cul}(e, t)$ and $\text{Hold}(e, t)$. The intuitive meaning of the first is that eventuality $e$ culminates at time $t$. The second one says that $e$ is either an eventuality in its development phase or that $e$ is a state. Given these assumptions a sentence like *Mary drew a circle* is translated into first-order formula (27), $S$ is again short for speech time.
The semantic effect of the progressive consists in transforming events into states. If ‘A’ is an event verb, then ‘be A-ing’ is to be treated semantically as a state verb; otherwise, ‘be A-ing’ is to be treated the same as ‘A’. Parsons [1989], p 222

With this assumption the logical representation of *Mary was drawing a circle* is (28).

Obviously (27) does not follow from (28). But (28) does imply that there is a circle Mary was drawing. Of course this circle may not be complete. This means that Parson’s ontology is bound to include incomplete objects. For a discussion of incomplete objects see for instance Baggio and van Lambalgen [2007].

According to the above quote the progressive doesn’t change anything in the case of activities. Therefore *Mary pushed a cart* and *Mary was pushing a cart* are predicted to be equivalent.

Parsons’ theory rests on the difference between the predicates *Cul* and *Hold*. This difference is left to intuition. No axiomatisation of these predicates is given. This is one of the objections in Zucchi [1999]. Zucchi also notes that under certain circumstances it is possible to derive a specific version of the imperfective paradox in Parsons’ theory. A careful discussion of Dowty’ and Parsons’ approaches and a combination of the two systems is contained in Landman [1992]. We will come back to the imperfective paradox in Section 7.

Having introduced events as primitive entities the philosophical question arises how events and time are related. A construction which shows that the time line can be constructed from events is due to Russell, Wiener and Kamp. We will sketch this construction in the next section.

### 6. From Events to Time

Russell remarked in [Russell, 1914, lecture IV]: “Even if there be a physical world such as the mathematical theory of motion presupposes, impressions on our sense-organs produce sensations which are not merely and strictly instantaneous, and therefore the objects of sense of which we are immediately conscious are not strictly instantaneous. Instants, therefore, are not among the data of experience and must be either inferred or constructed. It is difficult to see how they can be validly inferred; thus we are left with the alternative that they must be constructed.” We will present such a construction, as developed by Kamp [1979] who took up Russell’s ideas and slightly modified them. The construction takes event structures as primitives from which the structure of time, i.e. a total order, can be derived.

An event structure \( \langle E, P, O \rangle \) simply consists of a set of events \( E \) and the relations \( P \) (prcedes) and \( O \) (overlaps) and is characterized by axioms A1–7.
Definition 3. Let \( \langle E, P, O \rangle \) be a structure which satisfies axioms A1–7.

1. An instant of \( \langle E, P, O \rangle \) is a maximal subset of pairwise overlapping events, i.e. \( i \) is an instant of \( \langle E, P, O \rangle \) iff
   - (a) \( i \subseteq E \)
   - (b) For any \( a, b \in i : O(a, b) \)
   - (c) if \( e \in E \) but \( e \notin i \) then there is \( d \in i \) such that \( \neg O(d, e) \)

2. Let \( I \) be the set of instants of \( \langle E, P, O \rangle \)
3. For \( i_1, i_2 \in I \) put \( i_1 < i_2 \) iff there are \( e_1 \in i_1 \) and \( e_2 \in i_2 \) such that \( P(e_1, e_2) \)

We will first illustrate the Russell-Kamp construction by way of an example and then show that \( \langle I, \langle \rangle \) has the desired properties in theorem 1: it is a strict linear ordering which allows to construct time intervals.

Example 1. Let \( E = \{a, b, c, d, e\} \) and \( P \) be the set of ordered pairs \( \{(a, c), (a, d), (a, e), (b, e), (c, e)\} \). Given A7, the relation \( O \) consists of the set of pairs \( \{(a, b), (b, c), (b, d), (c, d), (d, e)\} \). \( E \) contains three maximal subsets of pairwise overlapping events, i.e. \( i_1 = \{a, b\} \), \( i_2 = \{b, c, d\} \) and \( i_3 = \{d, e\} \). Furthermore, \( i_1, i_2 \) and \( i_3 \) are linearly ordered \( i_1 < i_2 < i_3 \) because \( a \in i_1 \), \( c \in i_2 \) and \( P(a, c) \) as well as \( c \in i_2, e \in i_3 \) and \( P(c, e) \), respectively, and similarly for \( i_1 < i_3 \).

A proof of the following theorem can be found in [Kamp, 1979, p. 379].

Theorem 1. If an event structure \( \langle E, P, O \rangle \) satisfies A1–7, then the structure \( \langle I, \langle \rangle \) constructed according to definition 3 is a strict linear ordering, i.e. the following three conditions hold:

1. \( \forall i_1, i_2 (i_1 < i_2 \rightarrow \neg i_2 < i_1) \) (asymmetry)
2. \( \forall i_1, i_2, i_3 (i_1 < i_2 \land i_2 < i_3 \rightarrow i_1 < i_3) \) (transitivity)
3. \( \forall i_1, i_2 (i_1 \neq i_2 \rightarrow i_1 < i_2 \lor i_2 < i_1) \) (totality)

Once having constructed the set \( I \) it is possible to define time intervals and these, in turn, can be used to measure the temporal extent of an event. In example 1, for
instance, a goes on in the interval \([i_1]\), b in \([i_1, i_2]\), c in \([i_2]\), d in \([i_2, i_3]\) and e in \([i_3]\). Corollary 1 guarantees that these intervals can be constructed ([Kamp, 1979, p. 379]).

Corollary 1. For each \(e \in E\), the set \(\{i \in I | e \in i\}\) is a non-empty interval of \(\langle I, < \rangle\).

Proof: If \(i_1, i_2 \in e\) and \(i_1 < i < i_2\), it has to be shown that \(e \in i\), too. Suppose it is not. Then there is \(d \in i\) such that \(\neg O(d, e)\), hence \(P(d, e)\) or \(P(e, d)\). In the first case \(i < i_1\) and in the second case \(i_2 < i\), both yield a contradiction.

We cannot go into further details here, but refer the interested reader to Thomason [1986] who discusses how to construct time as a continuum – isomorphic to the real numbers – and uses a different construction originally proposed by Walker [1947]. Empirical evidence for events being ontological primitives will be presented in section 8 where we review psychological studies on event perception.

7. Event Calculus

In the following we will show that the imperfective paradox is an instance of the frame problem prominent in Artificial Intelligence (AI). To this end we will explain the event calculus which is an extension of McCarthy’s situation calculus (McCarthy [1977]) developed by Kowalski and Sergot [1986]. Then we will at least indicate how the event calculus allows for a formalisation of natural language tense and aspect by analysing the imperfective paradox.

Consider narrative (29).

(29) Yesterday morning Vincent loaded the gun. Then he smoked a cigarette. After smoking he pointed the gun at Fred and pulled the trigger.

Assuming that shooting at somebody with a loaded gun will lead to the death of that person we will interpret the discourse in a way that Fred gets eventually killed. Arriving at this inference is, however, not a trivial task. In (29) the frame problem is embodied in the form of the Yale Shooting Problem Hanks and McDermott [1986]. To demonstrate this, we will start with a very simple calculus – a variant of the Simple Event Calculus by Shanahan [1997] which only incorporates the very basic notion of instantaneous change. Later we will add axioms in order to deal with continuous change and then came back to the imperfective paradox. We will argue that the paradox can be viewed as one instance of the frame problem.

7.0.1. The Yale Shooting Scenario. Suppose we have three kinds of actions/events – Load, Smoke and Shoot – and three time dependent properties, the fluents Alive, Loaded and Dead\(^5\). Besides actions and fluents we will also need instants of time in our basic ontology. The predicates listed in Table 4 are used to express actions and their effects and to locate them in time.

Table 4: Predicates of the simple event calculus (from Shanahan, 1999)

\(^5\)The actions and fluents are assumed to be used uniquely. It is, for instance, not possible that Shoot and Smoke are identical actions. Technically, this can be done by adding a set of uniqueness-of-names-axioms (see e.g. Shanahan [1997]). For simplicity, we will leave them out.
The predicates are related to each other by a set of four simple axioms (30)-(33). The first three axioms state the conditions under which a fluent can hold at a time \( t \): either it holds right from the start or it is initiated at a time \( t_1 \) before \( t \) without a terminating action occurring in between the two. The fourth axiom defines the predicate \( \text{Clipped}(t_1, f, t_2) \) saying that a fluent \( f \) is clipped between times \( t_1 \) and \( t_2 \) if it is terminated by some action \( a \) happening between \( t_1 \) and \( t_2 \). Simplifying a whole lot, the axioms (30)-(33) formalize the notion of instantaneous change, such as two balls colliding or somebody dying. Moreover, they embody the notion of inertia.

In the following, all variables are assumed to be universally quantified with maximal scope. We will come to the semantics of the implication when we have introduced the complete scenario.

\[
\begin{align*}
(30) & \quad \text{Initially}(f) \rightarrow \text{HoldsAt}(f, 0) \\
(31) & \quad \text{Initially}(f) \land \neg \text{Clipped}(0, f, t) \rightarrow \text{HoldsAt}(f, t) \\
(32) & \quad \text{Happens}(a, t_1) \land \text{Initiates}(a, f, t_1) \land t_1 < t_2 \land \neg \text{Clipped}(t_1, f, t_2) \rightarrow \text{HoldsAt}(f, t_2) \\
(33) & \quad \text{Happens}(a, t) \land t_1 < t < t_2 \land \text{Terminates}(a, f, t) \rightarrow \text{Clipped}(t_1, f, t_2)
\end{align*}
\]

Next, we have to translate the discourse (29) into the Simple Event Calculus. The first set of formulas (34) – (36) states what the mentioned actions do. These formulas provide a crude sketch of the lexical meaning of \textit{load}, \textit{smoke} and \textit{shoot}. The effect of a \textit{Load} action is to make the fluent \textit{Loaded} hold, a \textit{Shoot} action makes \textit{Dead} hold (and \textit{Alive} not hold) with \textit{Loaded} being a precondition. Finally, \textit{Smoke} is assumed to have no effects. Thus, there are no rules involving \textit{Smoke}.

\[
\begin{align*}
(34) & \quad \text{Initiates}(\text{Load}, \text{Loaded}, t) \\
(35) & \quad \text{HoldsAt}(\text{Loaded}, t) \rightarrow \text{Initiates}(\text{Shoot}, \text{Dead}, t) \\
(36) & \quad \text{HoldsAt}(\text{Loaded}, t) \rightarrow \text{Terminates}(\text{Shoot}, \text{Alive}, t)
\end{align*}
\]

(29) comprises a \textit{Load} action followed by a \textit{Smoke} action followed by a \textit{Shoot} action. Using four arbitrarily chosen time points \( T_1 \)-\( T_4 \) the discourse can be represented as follows:

\[
\begin{align*}
(37) & \quad \text{Initially}(\text{Alive}) \\
(38) & \quad \text{Happens}(\text{Load}, T_1)
\end{align*}
\]

The predicates are related to each other by a set of four simple axioms (30)-(33). The first three axioms state the conditions under which a fluent can hold at a time \( t \): either it holds right from the start or it is initiated at a time \( t_1 \) before \( t \) without a terminating action occurring in between the two. The fourth axiom defines the predicate \( \text{Clipped}(t_1, f, t_2) \) saying that a fluent \( f \) is clipped between times \( t_1 \) and \( t_2 \) if it is terminated by some action \( a \) happening between \( t_1 \) and \( t_2 \). Simplifying a whole lot, the axioms (30)-(33) formalize the notion of instantaneous change, such as two balls colliding or somebody dying. Moreover, they embody the notion of inertia.

In the following, all variables are assumed to be universally quantified with maximal scope. We will come to the semantics of the implication when we have introduced the complete scenario.

\[
\begin{align*}
(30) & \quad \text{Initially}(f) \rightarrow \text{HoldsAt}(f, 0) \\
(31) & \quad \text{Initially}(f) \land \neg \text{Clipped}(0, f, t) \rightarrow \text{HoldsAt}(f, t) \\
(32) & \quad \text{Happens}(a, t_1) \land \text{Initiates}(a, f, t_1) \land t_1 < t_2 \land \neg \text{Clipped}(t_1, f, t_2) \rightarrow \text{HoldsAt}(f, t_2) \\
(33) & \quad \text{Happens}(a, t) \land t_1 < t < t_2 \land \text{Terminates}(a, f, t) \rightarrow \text{Clipped}(t_1, f, t_2)
\end{align*}
\]

Next, we have to translate the discourse (29) into the Simple Event Calculus. The first set of formulas (34) – (36) states what the mentioned actions do. These formulas provide a crude sketch of the lexical meaning of \textit{load}, \textit{smoke} and \textit{shoot}. The effect of a \textit{Load} action is to make the fluent \textit{Loaded} hold, a \textit{Shoot} action makes \textit{Dead} hold (and \textit{Alive} not hold) with \textit{Loaded} being a precondition. Finally, \textit{Smoke} is assumed to have no effects. Thus, there are no rules involving \textit{Smoke}.

\[
\begin{align*}
(34) & \quad \text{Initiates}(\text{Load}, \text{Loaded}, t) \\
(35) & \quad \text{HoldsAt}(\text{Loaded}, t) \rightarrow \text{Initiates}(\text{Shoot}, \text{Dead}, t) \\
(36) & \quad \text{HoldsAt}(\text{Loaded}, t) \rightarrow \text{Terminates}(\text{Shoot}, \text{Alive}, t)
\end{align*}
\]

(29) comprises a \textit{Load} action followed by a \textit{Smoke} action followed by a \textit{Shoot} action. Using four arbitrarily chosen time points \( T_1 \)-\( T_4 \) the discourse can be represented as follows:

\[
\begin{align*}
(37) & \quad \text{Initially}(\text{Alive}) \\
(38) & \quad \text{Happens}(\text{Load}, T_1)
\end{align*}
\]
Putting everything together, from the theory \((30)\wedge \ldots \wedge (41)\) we want to derive \(\text{HoldsAt}(\text{Dead}, T_4)\). Unfortunately, this sequent is not valid in the classical sense of being true in all models which satisfy the theory \((30)\wedge \ldots \wedge (41)\). Think for example of a situation in which the gun obscurely becomes unloaded while Vincent is smoking. In such a model, the precondition of the shooting action is clearly not met and thus the action will lack any effects. Note that nothing rules out such a model since it is entirely consistent with our theory \((30)\wedge \ldots \wedge (41)\). We thus have to further constrain the models under consideration. Crucially, we have to find a way to deal with the non-effects of actions; i.e. the frame problem.

How to proceed? From the fact that no terminating action was mentioned we want to conclude that no terminating event occurred. To do so, we have to strengthen the assumptions of the theory in a way that only those events are assumed to occur which have been explicitly stated in the discourse. The discourse model should be minimal, linking discourse understanding intimately to closed world reasoning. It is important to note that this strategy forces reasoning to be non-monotonic: adding further premises to a theory can make inferences invalid that have been been valid before.

**7.0.2. The completion of a program.** There are different techniques to formalize this line of reasoning; one is circumscription (e.g. Shanahan [1997]). Here, we will use the completion of a logic program adopting the Event Calculus of van Lambalgen and Hamm [2005]. Having a closer look at the formulas of the theory \((30)\wedge \ldots \wedge (41)\), an important feature to notice is that formulas come in two variants: they are either facts consisting of a simple predicate expression or they express rules with a conjunction of potentially negated formulas to the left of the implication sign and a single positive atomic expression to its right. In logic programming the former is called the body and the latter is called the head of a clause. We will now illustrate how the completion of a simple logic program is computed. Consider the description of a situation where the gun gets loaded at time 1 and a shooting event happens at time 10. This situation is stated in the following program:

\[
\begin{align*}
(42) & \quad \text{Happens} (\text{Load}, 1) \\
(43) & \quad \text{Happens} (\text{Shoot}, 10)
\end{align*}
\]

The uncompleted program does not yet rule out intervening events. It is, for instance, consistent with an \textit{Unload} event occurring at 9. The completion of the program should tell us that (42) and (43) were the only events. The completion is computed according to the following procedure\(^6\). We will start with the facts in (42) and (43). Both of these are \textit{Happens} formulae. The completion of the program

\(^6\)Precise definitions of logic programs and their completions can be found in Doets [1994] and Nienhuys-Cheng and de Wolf [1997].
intuitively corresponds to the fact that given this program $\text{Happens}(e, t)$ can only mean $\text{Load}$ to occur at 1 or $\text{Shoot}$ to occur at 10. We therefore first substitute variables for the constants and write the disjunction:

$$(44) \quad (e = \text{Load} \land t = 1) \lor (e = \text{Shoot} \land t = 10)$$

Then, we universally quantify over the variables $e$ and $t$ and strengthen the implication to a bi-implication:

$$(45) \quad \forall e \forall t (\text{Happens}(e, t) \leftrightarrow (e = \text{Load} \land t = 1) \lor (e = \text{Shoot} \land t = 10))$$

From (45) it follows that there were no intervening events. Strengthening the implications to bi-implications makes it impossible that a head can be true without the enabling conditions being met. The uniquely determined model of (45) is the minimal model of the logic program consisting of (42) and (43).

We can now come back to the Yale Shooting Problem in discourse (29). From the completion of the constraint logic program $((30)) \land \ldots \land ((41))$ it follows (non-monotonically) that $\text{HoldsAt} (\text{Dead}, T_4)$. Enriching the program with further information and computing the new completion, however, can lead to canceling this inference.

7.0.3. Introducing tense information. To serve as a discourse theory we need to incorporate tense. In van Lambalgen and Hamm [2005] this is done by adopting the classic tripartition into reference time, event time and speech time by Reichenbach (see section 3). When translating discourse (29) into Shanahan’s Event Calculus, we arbitrarily chose instants of time $T_1, \ldots, T_4$ to locate actions in time. This is not how it is done in discourse. The first sentence (yesterday morning Vincent loaded the gun) states that within some time interval (= reference time) before now (= speech time) there was a time $t'$ at which a loading event happened (= event time). The second sentence (Then he smoked a cigarette) is linked to this event by the connective then, stating that at some later time $t''$ a smoking event happened with $t''$ also being temporally located before now. Analogously for the third sentence: there is an event time $t'''$ at which $\text{Shoot}$ happens and $t'' < t''' < \text{now}$. Note that when locating events in time it is always done by existential statements of the form there is a time. So far, however, we have only been dealing with universally quantified variables. Clearly, a formula like $\forall t. \text{Happens}(\text{Load}, t) \land t < \text{now}$ cannot be used.

---

7To be complete we also have to show how to compute the completion of a rule. We will not go into this here. The interested reader is referred to [van Lambalgen and Hamm, 2005, definition 14]. To give an idea of how this works we will illustrate the completion of two propositional rules with the same head ($p_1, ..., p_n, r_1, ..., r_m, q$ all are propositions):

\(\text{(i)}\)

\(\begin{align*}
\text{a.} & \quad p_1 \land \ldots \land p_n \rightarrow q \\
\text{b.} & \quad r_1 \land \ldots \land r_m \rightarrow q
\end{align*}\)

The completion is:

\(\text{(ii)}\)

\(q \leftrightarrow (p_1 \land \ldots \land p_n) \lor (r_1 \land \ldots \land r_m)\)

This means that $q$ is false iff neither of the rules (i-a) or (i-b) is true.
to represent tense since it would state that \textit{Load} happens at all times in the past. We have to find a way how to introduce existentially quantified information.

Intuitively, it serves to introduce new entities into the domain of discourse. This explanation is at the heart of dynamic semantic theories such as Discourse Representation Theory (Kamp and Reyle [1993], see also the SEP entry for Discourse Representation Theory) or File Change Semantics (Heim [1982]) and We need an update procedure that introduces new actions/events into the discourse model but is compatible with the non-monotonic semantics introduced so far. Following van Lambalgen and Hamm [2005] we will use integrity constraints – a device from database theory (for a much more complete and formal treatment the reader is referred to Kowalski [1995] and [van Lambalgen and Hamm, 2005, ch. 8]) – to implement the notion of a minimal update of a discourse model. Consider discourse (46).

(46) The gun is loaded. Vincent is ready to shoot.

In interpreting the first sentence, we want to update the situation model in a way that makes the sentence true. The tense information will be interpreted as a goal to minimally update the situation model with a fluent \textit{Loaded} that holds now. In our framework, semantic interpretation is thus closely linked to planning in that finding an interpretation for a sentences requires coming up with a sequence of actions that makes the goal succeed:

(47) Make the query \textit{?HoldsAt}(\textit{Loaded, now}) succeed.

To resolve this integrity constraint we will use the axioms of the Simple Event Calculus plus world knowledge about the consequences of a load action stated in (48).

(48) \textit{Initiates}(\textit{Load, Loaded, t})

In logic programming, the sort of reasoning required to resolve (47) is carried out by a derivation procedure called \textit{resolution}. It starts with the formula that has to be made true in the discourse model. Resolution proceeds by identifying rules which have the query as their consequent and substituting the consequent with the antecedent conditions of the rule, making the antecedent formulas new queries themselves. The resolution stops when the query cannot be further resolved, that is, when a plan has been computed whose preconditions are all fulfilled given an appropriate update of the discourse model. To get an impression of how this works we will illustrate the resolution of (47). First, we have to compute the completion of the program and whenever possible substitute the variables with constants, in our case \textit{Loaded} and \textit{now}. In logic programming this substitution is done automatically via unification. Here is the completion of the \textit{HoldsAt} predicate which forms the head of two axioms – (31) and (32).

(49) \textit{HoldsAt}(\textit{Loaded, now}) \leftrightarrow
    (\textit{Initially}(\textit{Loaded}) \land \lnot \textit{Clipped}(0, \textit{Loaded, now})) \lor
\[(\text{Happens}(a, s) \land \text{Initiates}(a, \text{Loaded}, s) \land s < \text{now} \land \\
\neg \text{Clipped}(s, \text{Loaded}, \text{now}))\]

Since the theory doesn’t contain the statement \(\text{Initially}(\text{Loaded})\), closed world reasoning yields that at time 0 the gun is not loaded \((\neg \text{Initially}(\text{Loaded}))\). No matter how we update the discourse representation the first disjunct can never succeed. We thus have to move on to the second disjunct. In this case, the database is searched for an action \(a\) and a time \(s\) such that \(\text{Initiates}(a, f, s)\), \(\text{Happens}(a, s)\) and \(\neg \text{Clipped}(s, \text{Loaded}, t)\). On the sole basis of the information provided in the discourse the subquery \(\text{Happens}(a, s)\) will fail, because in discourse (46) there is no information about a load action. However, since integrity constraints are intended to be made true, the database will be updated with a clause \(\text{Happens}(\text{Load}, s) \land s < \text{now}\) which makes the second disjunct true and the query succeed.

7.0.4. The imperfective paradox reconsidered. In the following, we will analyze the imperfective paradox as an instance of the frame problem (cf. Stenning and van Lambalgen [2005], van Lambalgen and Hamm [2005], Stenning and van Lambalgen [2008], Baggio and van Lambalgen [2007], Baggio et al. [2008]). Both activities and accomplishments involve continuous change and we therefore have to add additional predicates and axioms to properly deal with gradually changing objects which are under the influence of an external force. Table 7.0.4 introduces two new predicates; axioms A1–A5 provide a general theory of instantaneous and continuous change (from [van Lambalgen and Hamm, 2005, p. 40]). As can be easily seen, the axiom system extends that of Shanahan [1997].

<table>
<thead>
<tr>
<th>Formula</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\text{Releases}(e, f, t))</td>
<td>Fluent (f) starts to change due to event (e) at time (t)</td>
</tr>
<tr>
<td>(\text{Trajectory}(f_1, t, f_2, d))</td>
<td>If fluent (f_1) holds from (t) until (t+d), then (f_2) holds at (t+d)</td>
</tr>
</tbody>
</table>

A1) \(\text{Initially}(f) \rightarrow \text{HoldsAt}(f, 0)\)

A2) \(\text{HoldsAt}(f, r) \land r < t \land \neg \exists s < r \text{HoldsAt}(f, s) \land \neg \text{Clipped}(r, f, t) \rightarrow \text{HoldsAt}(f, t)\)

A3) \(\text{Happens}(e, t) \land \text{Initiates}(e, f, t) \land t < t' \land \neg \text{Clipped}(t, f, t') \rightarrow \text{HoldsAt}(f, t')\)

A4) \(\text{Happens}(e, t) \land \text{Initiates}(e, f_1, t) \land t < t' \land t' = t + d \land \text{Trajectory}(f_1, t, f_2, d) \land \neg \text{Clipped}(t, F_1, t') \rightarrow \text{HoldsAt}(f, t')\)

A5) \(\text{Happens}(e, s) \land t < s < t' \land (\text{Terminates}(e, f, s) \lor \text{Releases}(e, f, s)) \rightarrow \text{Clipped}(t, f, t')\)

Axioms 4 and 5 define continuous change. Axiom 4 defines the \(\text{Trajectory}\) predicate. To see what it says, let’s consider a situation of running a mile where \(f_1\) is instantiated by running and \(f_2\) by distance(x). Should running be true during the whole interval from \(t\) until \(t'\) then distance(\(a + 1\)) will be true at \(t'\). The value of \(x\), i.e. the actual distance at each time instant will be determined by the law of the process under consideration (e.g. running speed). Axiom 5 defines the predicate
Clipped\((t,f,t')\) which is true if \(f\) is terminated or released by an event \(e\) which happens in the time interval between \(t\) and \(t'\). Thus Clipped covers both instantaneous and continuous change. The axioms provide a general theory of what can change and what stays constant. For concrete situations we also need the specific temporal and causal relationships. This kind of information is specified in so called scenarios representing lexical meaning. The lexical meaning of the accomplishment run a mile corresponds to scenario (50). Like the axioms, the scenario takes the form of a logic program. All variables are universally quantified.

\[
\begin{align*}
(50) & \quad a. \text{Initially}(distance(a)) \\
& \quad b. \text{Initiates}(start, run, t) \\
& \quad c. \text{Releases}(start, distance(x)) \\
& \quad d. \text{Initiates}(finish, distance(a + 1), t) \\
& \quad e. \text{Terminates}(finish, run, t) \\
& \quad f. \text{HoldsAt}(run, t) \land \text{HoldsAt}(distance(a + 1), t) \rightarrow \text{Happens}(finish, t) \\
& \quad g. \text{HoldsAt}(distance(x), t) \rightarrow \text{Trajectory}(run, t, distance(x + g(d)), d)
\end{align*}
\]

Scenarios formalize the event nucleus of Moens and Steedman [1988]. Every accomplishment takes the same form of scenario with the only difference that the individual scenarios involve different preparatory processes (run, build, etc.), incremental themes (distance(x), house(x), etc.) and resultant states. Furthermore, it is obvious that the preparatory process run, an activity, is a proper part of scenario (50). The simpler activity scenario can be easily arrived at if we remove clauses (d)–(f) from (50). We are now in the position to solve the imperfective paradox avoiding the problems discussed in section 5. Both, a past progressive activity sentence as in (51-a) and a past progressive accomplishment sentence (51-b) will trigger a discourse update as stated in integrity constraint (51-c).

\[
\begin{align*}
(51) & \quad a. \text{John was running.} \\
& \quad b. \text{John was running a mile.} \\
& \quad c. \ ?\text{HoldsAt}(run, t) \land t < now \text{ succeeds.}
\end{align*}
\]

The discourse update of the completed program (50) yields that at some point after \(t\) John will achieve the culmination and will have run one mile (see [van Lambalgen and Hamm, 2005, p. 61ff.] for a proof). Obviously, the same integrity constraint in combination with the activity in (51-a) only licenses the inferences that there was some past running activity. In the case of an accomplishment the course of events dramatically changes if we add information to the representation as in (52-b). Now, both (52-a) and (52-b) only allow the inference that there was a running event but, arguably, John didn’t reach his goal in (52-b). Without going into further details it should be clear how this non-monotonic inference works. The derivation works analogously to the simpler instances of the frame problem discussed above.

\[
\begin{align*}
(52) & \quad a. \text{John was running, when he was hit by a truck.} \\
& \quad b. \text{John was running a mile, when he was hit by a truck.}
\end{align*}
\]
The just outlined solution to the imperfective paradox makes reference to the intentions of the agent involved in the event. Therefore the notion of a goal or intention is built into the system right from the start. However, it does not involve possible worlds with problematic notions like inertia worlds but based upon minimal models and non-monotonic reasoning instead. This has additional advantages. Firstly, the Event Calculus will compute a minimal model in case an integrity constraint can be satisfied and the number of construction steps yields a precise complexity measure which, for instance, allows us to derive predictions for cognitive processing. Secondly, the non-monotonic nature of the Event Calculus can be used to model the incremental construction of a temporal model with inferences which hold locally at a discourse segment $s_n$ but which can be undone at some later discourse unit $s_{n+1}$. This is a necessary prerequisite if we think of interpreting sentences like (52-b) from left to right assuming incremental interpretation (see e.g. Baggio and van Lambalgen [2007] and Bott [2010] for implementations of the Event Calculus as a processing model). In the next section we will review psychological and psycholinguistic work on the interpretation of aspect which will provide empirical motivation for the outlined analysis.

8. COGNITIVE ASPECTS

For limitations of space, we will not go into psychological and psycholinguistic studies on tense but instead focus on processing studies on aspect and within this topic limit ourselves to studies of the adult system. The interested reader is referred to Dickey [2001] and the references therein for psycholinguistic work on tense as well as to the SEP entry The Experience and Perception of Time and the contributions in Grondin [2008] for psychology of time in general.

Psychologists have recently turned to study event perception and its neural basis (e.g. Zacks and Tversky [2001], Zacks et al. [2001], Zacks et al. [2006]). The underlying question behind this line of research is whether events play a role in how we cognitively structure and remember changing states of affairs and how they are represented. Zacks et al. [2001] use an analogy from the domain of objects to the domain of events to define an event as a “segment in time at a given location that is conceived by an observer to have a beginning and an end” (p.30). They claim that “the ability to identify the parts of events and their relationships constitutes a distinct perceptual process” (p.30) which they call event structure perception. Just like an object is an ontological primitive in the spatial domain an event is an ontological primitive in the spatio-temporal domain\(^8\). Objects are recognized by shape, color etc. and have boundaries in space. Analogously, events have boundaries in time, but are also bounded in space. An event of buttering a toast, for instance, happens at a particular time, but also in a particular location in space and is therefore spatially bounded. Continuing the analogy, both objects and events can be identified and categorized using hierarchical relations.

In Zacks et al.’s experiments, participants segmented an ongoing activity (like washing dishes) while watching it on film by pressing a key to mark “natural and

\(^8\)This agrees with Vendler’s philosophical position in chapter five of Vendler [1967].
meaningful” unit boundaries. The grain at which participants segmented the activity was manipulated between subjects: one group was asked to mark the largest meaningful units (coarse grain size), the other group the smallest units (fine grain size) which still could be considered to be complete events. The placement of perceived event boundaries provides information about the psychological status of events and their partonomic hierarchy. If the stream of action is perceived as consisting of discrete events participants should place event boundaries consistently and this is what Zacks et al. [2001] found. Further, if participants make use of partonomic hierarchies coarse event boundaries should be aligned with fine boundaries, since the end of each superordinate event is also the end of its last subevent. Again, this hypothesis was corroborated by a strong hierarchical bias effect. Events thus seem to be psychologically real and to be hierarchically structured.

8.1. Lexical Aspect. During the last decade psycholinguists have started to investigate the cognitive reality of lexical aspect and how it is processed. The existing studies can be sub-classified into studies on Aktionsart simpliciter and shifts from one aspectual class to an other, i.e. aspectual coercion. In the following, we will provide a coarse overview over both areas of research.

McKoon and Macfarland [2002] were among the first to study processing consequences of decompositional analyses along the lines of Dowty [1979]. They provided evidence from reading times and lexical decision times that accomplishments (α CAUSE x BECOME IN-STATE) are inherently more complex than achievements (x BECOME IN-STATE). Similarly, Gennari and Poeppel [2003] compared eventive predicates (which included accomplishments, achievements and activities) to statives and found that the former were more complex to process than the latter. Brennan and Pylkkänen [2010] extended this line of research to psychological verbs and compared accomplishments (e.g. scare) with statives (e.g. cherish) using reading time methods and magnetoencephalography (MEG). They also included a comparison of the simple psychological statives with modified sentences that required coercion (e.g. within half an hour, the child cherished the precious kitten). Reading times and MEG data indicated that accomplishments were more complex to comprehend than statives and, moreover, that the enhanced semantic complexity led to different MEG components than aspectual coercion. Finally, Coll-Florit and Gennari [2011] compared durative states and punctual event predicates (mostly achievements) and observed longer reading times of statives than of event predicates. They attributed this effect - which runs counter to what would be expected under a decomposition analysis - to the fact that durative situations occur in semantically more diverse contexts and elicit more diverse associations than event predicates which may modulate the required processing effort. Taken together, the existing studies provide evidence for complexity differences between the aspectual classes lending support to decompositional analyses. However, not all issues are entirely solved yet and further research is required.

The second line of research concerns aspectual coercion. Here, the main research question has been whether aspectual coercion is a costly operation. Existing research has thereby almost exclusively focused on one coercion type, i.e. coercion
of point action verbs into an iterative interpretation (but see Brennan and Pylkkänen [2010], Bott [2010] for exceptions). Again, the findings are mixed. Early studies have provided evidence for coercion costs employing secondary tasks such as cross-modal lexical decision or stop making sense judgments (Piñango et al. [1999], Todorova et al. [2000]). Pickering et al. [2006] used the same materials as in the experiments mentioned above, but tested a coerced meaning during ordinary reading without an additional task. In two self-paced reading and two eyetracking experiments, they found aspectual coercion to be no more difficult than their aspectual control conditions. This lack of effect let them propose the aspectual underspecification hypothesis, stating that the aspectual representation stays underspecified during normal reading. Brennan and Pylkkänen [2008] challenged this view and reported a coercion effect of coercion sentences like (53-a) as compared to aspectual controls (53-b) both in self-paced reading and in MEG; on the basis of a rating study they had carefully selected clear instances of point action verbs. Their MEG study revealed activation in the anterior midline field, a MEG component that has been observed for other non-aspectual cases of coercion, too.

(53)  

a. Throughout the day, the student sneezed in the back of the classroom.  
b. After twenty minutes, the student sneezed in the back of the classroom.

Complicating matters, Bott [2008, 2010] applied the same norming procedures, but did not find evidence of coercion cost in iterative point action verbs in German. With other types of aspectual coercion, however, there were clear indications of processing difficulty indicating that the underlying processes differ between coercion types. Furthermore, the study provided evidence that at least in some cases (achievement → accomplishment) coercion operations are not triggered by an aspectual mismatch, but can proceed smoothly by enriching the aspectual representation with additional eventualities. This was shown in an event-related potentials (ERP) study in which we observed a double dissociation of ERP components elicited by sentences with aspectual coercion as compared to sentences that contained an aspectual mismatch.

8.2. Grammatical Aspect. Grammatical aspect has recently received increased interest in psycholinguistics (for a comprehensive review see Madden and Ferretti [2009]). An interesting line of studies has investigated (e.g. Ferretti et al. [2007]) the accessibility of event participants in English progressive and simple sentences. They provide evidence that in line with the linguistic description provided above the progressive presents an event from the inside making participants, instruments and places fully accessible, whereas the simple forms present events as complete units and the event participants are less accessible. Using picture selection/verification Madden and Zwaan [2003] found that participants were faster and more likely to choose a picture showing a complete event rather than a picture depicting an ongoing event after they had read a sentence with an accomplishment verb in the simple past. This indicates that English speakers encode accomplishments in the simple past as complete events. This finding was supported by a study by Anderson et al.
They compared accomplishments describing a path such as *Tom jogged/was jogging to the woods and then stretched when he got there*. Participants listened to these sentences while they had to use the computer mouse to drag and drop a human character in a visual scene. When hearing a sentence in the past progressive many drops took place at the beginning and the center of the path, whereas in the simple past most drops were at the end of the path. Interestingly, the differences were by no means categorical. Even in the simple past condition the character was in some trials positioned well before the end of the path. This indicates that accomplishments in the simple past are consistent with incomplete events. Another study relevant for the discussion in the previous sections is an ERP study on the imperfective paradox by Baggio et al. [2008]. They investigated the processing of Dutch imperfective sentences such as *het meisje was een brief aan het schrijven* (*the girl was writing a letter*) which were either followed by a *when*-clause that made the culmination unlikely (when her friend spilled coffee on the paper) or by a sentence which was compatible with the attainment of the goal (when her friend spilled coffee on the tablecloth). The first kind of sentences elicited larger sustained anterior negativities compared to sentences that were compatible with the culmination. Interestingly, the amplitude of the negativity was correlated with the frequency with which participants responded that the culmination was not reached. These findings indicate that a progressive sentence triggers a default inference to a complete event which can be canceled again, if the context so requires.

Grammatical aspect is subject to clear crosslinguistic differences. This raises the question whether the grammatical system has an influence on how we process language. In von Stutterheim et al. [2009] the authors reported findings of a production study comparing event descriptions of English, German and Dutch speakers that were elicited while they were watching and reporting a silent video clip. Their study revealed clear differences between productions from the three languages. English speakers used the progressive to start event descriptions well before the endpoint was visible (e.g. *A car is going down a lane... to a farmhouse*). Germans showed a different behavior relating their descriptions to the endpoint of a motion event. It thus seems as if production is constrained by the language in which it is realized. Recently, von Stutterheim et al. [2012] extended this line of research to a sample of seven languages and showed that the aspectual properties of a language influence how speakers conceptualize events. It’s not only that speakers of different languages talked differently about motion events (e.g. mention of endpoints), their language also had a clear influence on their looking behavior while watching the movies and influenced their memory capacity for those parts of the scene corresponding to endpoints. As for comprehension, Bott and Hamm [submitted] investigated how the aspectual system of a language influences processing difficulty. They compared coercion of German and English accomplishments which were modified by *for x time* adverbials. (54) is a sample item.

(54) a. Der Architekt errichtete das Monument zwei Jahre lang, aber dann wurde ihm der Kredit gekündigt.
b. The architect built the monument for two years, but then the bank canceled his credit.

In the German experiments, condition (54-a) was read as fast as a non-coercing control condition with an in x time adverbial. This was different in English where the authors found coercion to cause a substantial slowdown of reading speed. They interpret this crosslinguistic difference as follows: languages which have the grammatical means to express an aspectual difference via alternative forms - the progressive vs. simple forms in English - enforce immediate aspectual commitment, whereas languages that lack this grammatical feature have to leave it underspecified. In other words, German readers will leave it to the context whether an accomplishment expresses a complete or incomplete event, whereas English readers immediately strengthen an accomplishment in the simple form into a complete event interpretation. They take this strengthening to be a pragmatic process that is due to competition between alternative grammatical forms.

To summarize, the existing studies lend empirical support to the aspectual distinctions made in the semantic and philosophic literature. Moreover, the grammatical system of a language influences the way language is processed and may even influence the way we experience the world.

9. CONCLUSION

We introduced and discussed several important semantic, philosophical and technical concepts and theories of temporality and at least indicated how these concepts and theories are related to cognition. However, we did this exclusively from a sentence internal perspective. In their unfortunately still unpublished book Kamp and Rohrer [1986] showed that tense and aspect are important means for discourse organization as well. As outlook we present their French examples and explain the role tense and aspect play for discourse organization. Consider sentence (55) in which the passé simple (PS) occurs four times.

(55) Pierre se leva, monta dans sa chambre, ferma la porte et alluma la radio.

First the PS in (55) provides temporal information; all events are located in the past. But the PS conveys aspectual information too. It says that the internal constitution of the events is not important and this means that PS expresses perfectivity. PS imposes a view of the events “from the outside”. This is then taken to explain why multiple uses of the PS imply a succession of the events described as witnessed in example (55). This means that Pierre getting up precedes his going up to his room and this events precedes closing the door which finally precedes turning the radio on. Thus, the PS structures the above little discourse in a particular way. Early discourse representation theory explains this effect by assuming that the PS introduces a new reference point after a previously introduced one.

---

9See also Kamp [1991].

10This is certainly to coarse; see Kamp and Rohrer [1986] for a more detailed account and also for some counterexamples.
The discourse function of the French imparfait is different. Example (56-a) shows that like the past in English the imparfait\(^\text{11}\) is not felicitous when uttered out of the blue.

(56) a. *Il faisait chaud.

Kamp explains these facts in the following way. The imparfait does not introduce its own reference time and in (56-a) there is no previous reference point given. Therefore (56-a) is out. In (56-b) the sentence in the passé simple Jean ôta sa veste introduces a reference time that can be used to anchor *Il faisait chaud. The imperfective aspect of the imparfait is then explained by the fact that the PS event happens while the sentence in the imparfait holds. Thus passé simple and imparfait not only provide temporal information but they also serve as means to structure discourse in different ways.

Of course these brief remarks about the discourse function of tense and aspect hardly touch the tip of the iceberg. For more information on the discourse semantics of French tense and discourse organization in general the reader is advised to consult de Swart and Corblin [2002], Asher and Lascarides [2003] and van Lambalgen and Hamm [2005], especially chapter 9. This short note on discourse structure completes our article on tense and aspect.

REFERENCES


O. Bott and F. Hamm. Crosslinguistic variation in the processing of aspect. In B. Hemforth, B. Schmiedtová, and C. Fabricius-Hansen, editors, Experimental

\(^{11}\)See our discussion of example (8-a) in section 3.


