Table of contents

List of tables xi
List of figures xiii
Abbreviations and symbols xv
Acknowledgments xvii

CHAPTER 1
Introduction 1

CHAPTER 2
Events in cognitive psychology and linguistics 3
  2.1 Events from a psychological perspective 3
     2.1.1 Experiments on event perception and reporting 5
     2.1.2 The importance of planning in narrating
          and remembering events 7
     2.1.3 A brief summary 8
  2.2 Events in linguistics 9
     2.2.1 A calculus of events (Hamm & van Lambalgen (2005)) 11
     2.2.2 The representation of Aktionsart 16
     2.2.3 The classification of Aktionsart 24
     2.2.4 The composition of Aktionsart 26
     2.2.5 Hierarchical event structures 29
     2.2.6 A glimpse at the computational machinery 31
  2.3 Coercion 33
     2.3.1 Semantic theories on coercion 34
     2.3.2 Operator-based accounts 38
     2.3.3 Underspecification 42
     2.3.4 Planning accounts 47
     2.3.5 A short summary 50
     2.3.6 Psycholinguistic studies on coercion 50
     2.3.7 Studies on complement coercion 51
     2.3.8 Studies on aspectual coercion 54
  2.4 Summary and conclusions 59
CHAPTER 7
The processing of temporality in the brain 193

7.1 EEG research in psycholinguistics 193
   7.1.1 Measuring event-related brain potentials 193
   7.1.2 Neurolinguistic models of sentence comprehension 200

7.2 Relevant brain studies 204
   7.2.1 Brain studies on temporality in language 204
   7.2.2 Brain studies on complement coercion 207

7.3 An EEG study on the processing of Aktionsart
   and tense (Experiment 9) 208
   7.3.1 Theoretical accounts of aspectual coercion 209
   7.3.2 Method 216
   7.3.3 Results 220
   7.3.4 Discussion 224

7.4 Summary and conclusions 230
   7.4.1 A question for future research 232

CHAPTER 8
The specified processing model of aspectual reanalysis 235

8.1 A replication of Experiments 5 and 6: Experiment 10a/b 238
   8.1.1 Materials 239
   8.1.2 Pretesting the readings 240

8.2 Iterative readings: Experiment 10a and b 242
   8.2.1 Method 242
   8.2.2 Results 243
   8.2.3 Discussion 245

8.3 Summary and conclusions 248

CHAPTER 9
Summary and conclusions 251

9.1 Relating the findings to formal semantic accounts
   of aspectual reinterpretation 251
   9.1.1 Operator-based accounts 251
   9.1.2 Underspecification accounts 252
   9.1.3 Planning accounts 254

9.2 Predictions for psycholinguistic experiments 257

9.3 Relating the findings to psycholinguistic studies on coercion 258
   9.3.1 Studies on aspectual coercion 258
   9.3.2 Studies on complement coercion 260

9.4 Relating the findings to models of sentence processing 261
   9.4.1 Towards an immediacy model of semantic processing 262

9.5 Open questions 265
   9.5.1 Questions for further research 265
   9.5.2 Issues in modeling aspectual processing 266

References 269

Appendix
A. The semantics of logic programming and its implementation
   by recurrent neural nets 279
   A.1 Propositional logic programming 279
   A.2 The construction of minimal models using neural nets 292
B. Discourse representation theory (DRT) 289
   B.1 Processing a sample discourse 290
C. Target sentences in experiment 1 295
D. Lexical frequencies of verbs in Experiment 2 299
E. Target sentences in Experiment 2 301
F. Discourses in Experiment 3 307
G. Target sentences in Experiment 4 (4a and 4b) 313
H. Target sentences in Experiment 5 323
I. Target sentences in Experiment 6 327
J. Target sentences in Experiment 8 333
K. Target sentences in Experiment 9 343
L. Normed fillers 367
M. Target sentences in Experiment 10a 369
N. Target sentences in experiment 10b 377
# List of tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Aspectual tests</td>
<td>25</td>
</tr>
<tr>
<td>2.2</td>
<td>The standard Aktionsarten and their feature distribution</td>
<td>28</td>
</tr>
<tr>
<td>3.1</td>
<td>The compositional construction of the aspects</td>
<td>78</td>
</tr>
<tr>
<td>4.1</td>
<td>Sample frequency distribution for <em>backen</em> and <em>beladen</em></td>
<td>116</td>
</tr>
<tr>
<td>4.2</td>
<td>Judgments in Experiment 2</td>
<td>121</td>
</tr>
<tr>
<td>4.3</td>
<td>Comparisons between &quot;yes, makes sense&quot; answers in Experiment 2</td>
<td>122</td>
</tr>
<tr>
<td>4.4</td>
<td>Reading times on the adverbial in Experiment 2</td>
<td>123</td>
</tr>
<tr>
<td>4.5</td>
<td>Reading times on the two adverbial regions in Experiment 3</td>
<td>131</td>
</tr>
<tr>
<td>5.1</td>
<td>Reading times in Experiment 4a conditionalized on judgments</td>
<td>143</td>
</tr>
<tr>
<td>5.2</td>
<td>End of sentence reading times in Experiment 6</td>
<td>163</td>
</tr>
<tr>
<td>6.1</td>
<td>Distribution of different completion types in Experiment 7</td>
<td>177</td>
</tr>
<tr>
<td>6.2</td>
<td>Judgment times in Experiment 7</td>
<td>187</td>
</tr>
<tr>
<td>7.1</td>
<td>Time lapse for sample trials in Experiment 9</td>
<td>217</td>
</tr>
<tr>
<td>7.2</td>
<td>Mean amplitude in voltage for coercion compared to control in late time window 900–1500 ms post participle onset</td>
<td>227</td>
</tr>
<tr>
<td>8.1</td>
<td>Hypotheses that have been confirmed (= ✓) or disconfirmed (= No)</td>
<td>236</td>
</tr>
<tr>
<td>c.1</td>
<td>Items used in Experiment 1</td>
<td>295</td>
</tr>
<tr>
<td>d.1</td>
<td>Frequency distribution of verbs used in Experiment 2</td>
<td>299</td>
</tr>
<tr>
<td>e.1</td>
<td>Items used in Experiment 2</td>
<td>301</td>
</tr>
<tr>
<td>f.1</td>
<td>Items used in Experiment 3</td>
<td>307</td>
</tr>
<tr>
<td>g.1</td>
<td>Items used in Experiment 4</td>
<td>313</td>
</tr>
<tr>
<td>h.1</td>
<td>Items used in Experiment 5</td>
<td>323</td>
</tr>
<tr>
<td>i.1</td>
<td>Items used in Experiment 6</td>
<td>327</td>
</tr>
<tr>
<td>j.1</td>
<td>Items used in Experiment 8</td>
<td>333</td>
</tr>
<tr>
<td>k.1</td>
<td>&quot;Aktionsart&quot;-Items</td>
<td>343</td>
</tr>
<tr>
<td>k.2</td>
<td>&quot;Tense&quot;-Items</td>
<td>368</td>
</tr>
<tr>
<td>m.1</td>
<td>Items used in Experiment 10a</td>
<td>369</td>
</tr>
<tr>
<td>n.1</td>
<td>Items used in Experiment 10b</td>
<td>377</td>
</tr>
</tbody>
</table>
List of figures

2.1. Graphic illustration of the semantic properties of the verbs walk, reach & know 27
2.2. The aspectual transition network (from Steedman, 1997) 37
3.1 A model for aspectual interpretation and reinterpretation 62
3.2 Cumulated rejection rates in Exp. 1 95
3.3 Reading times for “yes, makes sense”-responses in Exp. 1 96
4.1 The verbs ordered according to their difference scores 117
4.2 Mean percent of activity and accomplishment uses in dependence of the two aspectual classes plus 95% confidence intervals 118
4.3 Mean adverbial effects in milliseconds 122
4.4 Sensicality judgments in Experiment 3 129
4.5 Mean reading times per character in Experiment 3 130
5.1 Distribution of judgments over participants 141
5.2 Reading times phrase by phrase in Experiment 4a. 142
5.3 Judgment data in Experiment 5. 148
5.4 Mean reading times per character in Experiment 5 in the VOA order on the left-hand side and in the VAO order on the right-hand side. 149
5.5 Mean reading times per character in Experiment 5 for activities. 150
5.6 Grammaticality ratings of the experimental sentences. 160
5.7 Mean reading times in Experiment 6. 162
6.1 Mean grammaticality judgments of the three word order variants. 171
6.2 Percent “nonsense” answers in Experiment 7. 176
6.3 Sensicality judgments of Experiment 4a and of Experiment 4b. 181
6.4 Reading times per character in Experiment 4b. 182
6.5 Mean of “yes, makes sense” judgments in Experiment 8. 186
6.6 Mean reading times in Experiment 8. 188
7.1 Sample ERP experiment. 194
7.2 Semantic acceptability judgments in the ERP study. 220
7.3 Grand averages elicited by aspectual mismatch, coercion and control. 221
7.4 P600 due to aspectual mismatch. 222
7.5 Working memory LAN due to coercion. 223
7.6 Mean amplitude for the three aspectual conditions 500–900 ms post stimulus with respect to anteriority and hemisphere. 224
7.7 Grand averages of tense violations compared to control. 225
7.8 P600 due to a tense violation. 226
8.1 Mean judgments in the pretest of Experiments 10a and b. 241
8.2 Mean sensicality judgments in Experiments 10a and b. 243
8.3 Mean reading times in Experiment 10a. 244
8.4 Mean reading times in Experiment 10b. 245
8.5 The specified model of aspectual processing. 249
A.1 Truth tables for three-valued connectives. 280
A.2 Representation of the node corresponding to a proposition p. 283
A.3 Three valued negation. 284
A.4 Recurrent net representing the program [p]. 285
A.5 Recurrent net representing the program [p; ¬ab; p ∧ ¬ab → q]. 286
A.6 Recurrent net representing the program [p; r → ab; p ∧ ¬ab → q]. 286

### Abbreviations and symbols

<table>
<thead>
<tr>
<th>Symbol/Phrase</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>#</td>
<td>semantically deviant sentence</td>
</tr>
<tr>
<td>*</td>
<td>clearly unacceptable sentence</td>
</tr>
<tr>
<td>?</td>
<td>marked sentence</td>
</tr>
<tr>
<td>[± ADDTO]</td>
<td>feature expressing whether a verb introduces a path or not</td>
</tr>
<tr>
<td>[± DUR]</td>
<td>feature expressing whether a verb is durative or not</td>
</tr>
<tr>
<td>[± SQA]</td>
<td>specified quantity feature</td>
</tr>
<tr>
<td>[± PERF]</td>
<td>perfectivity feature</td>
</tr>
<tr>
<td>α</td>
<td>conventional significance level</td>
</tr>
<tr>
<td>ξ</td>
<td>existential quantifier</td>
</tr>
<tr>
<td>∀</td>
<td>universal quantifier</td>
</tr>
<tr>
<td>λ</td>
<td>lambda operator</td>
</tr>
<tr>
<td>¬</td>
<td>negation operator</td>
</tr>
<tr>
<td>→</td>
<td>conditional</td>
</tr>
<tr>
<td>↔</td>
<td>biconditional</td>
</tr>
<tr>
<td>∨</td>
<td>disjunction</td>
</tr>
<tr>
<td>∧</td>
<td>conjunction</td>
</tr>
<tr>
<td>acc.</td>
<td>accomplishment</td>
</tr>
<tr>
<td>ach.</td>
<td>achievement</td>
</tr>
<tr>
<td>act.</td>
<td>activity</td>
</tr>
<tr>
<td>AdvP</td>
<td>adverbial phrase</td>
</tr>
<tr>
<td>AEH</td>
<td>aspectual enrichment</td>
</tr>
<tr>
<td>AKTIONSART</td>
<td>aktionsart feature</td>
</tr>
<tr>
<td>AMF</td>
<td>anterior midline field, a MEG component</td>
</tr>
<tr>
<td>ANOVA</td>
<td>(repeated measures) analysis of variance</td>
</tr>
<tr>
<td>AP</td>
<td>adjective phrase</td>
</tr>
<tr>
<td>ATSH</td>
<td>abstract type shifting hypothesis</td>
</tr>
<tr>
<td>AUH</td>
<td>aspectual underspecification hypothesis</td>
</tr>
<tr>
<td>CI</td>
<td>confidence interval</td>
</tr>
<tr>
<td>CLLS</td>
<td>constraint language for lambda structures</td>
</tr>
<tr>
<td>CP</td>
<td>complementizer phrase</td>
</tr>
<tr>
<td>CVPH</td>
<td>complete verb phrase</td>
</tr>
<tr>
<td>DP</td>
<td>determiner phrase</td>
</tr>
<tr>
<td>DRS</td>
<td>discourse representation structure</td>
</tr>
<tr>
<td>DRF</td>
<td>discourse representation theory</td>
</tr>
<tr>
<td>DRF/EC</td>
<td>semantic framework combining DRT and EC</td>
</tr>
<tr>
<td>e_0, e_1</td>
<td>events</td>
</tr>
<tr>
<td>EC</td>
<td>event calculus</td>
</tr>
<tr>
<td>EEG</td>
<td>electroencephalography</td>
</tr>
<tr>
<td>ELAN</td>
<td>early left anterior negativity in the ERPs</td>
</tr>
<tr>
<td>ERP</td>
<td>event related potentials</td>
</tr>
<tr>
<td>F_1</td>
<td>F value of the by subjects ANOVA</td>
</tr>
<tr>
<td>F_2</td>
<td>F value of the by items ANOVA</td>
</tr>
<tr>
<td>f_0, f_1</td>
<td>fluencts</td>
</tr>
<tr>
<td>fMRI</td>
<td>functional magnetic resonance imaging</td>
</tr>
<tr>
<td>IAIH</td>
<td>incremental aspectual interpretation hypothesis</td>
</tr>
<tr>
<td>IP</td>
<td>inflection phrase</td>
</tr>
<tr>
<td>KI</td>
<td>kilo ohm</td>
</tr>
<tr>
<td>LAIH</td>
<td>late aspectual interpretation hypothesis</td>
</tr>
<tr>
<td>LAN</td>
<td>left anterior negativity in the ERPs</td>
</tr>
<tr>
<td>LDH</td>
<td>lexical disambiguation hypothesis</td>
</tr>
<tr>
<td>LPH</td>
<td>lazy parsing hypothesis</td>
</tr>
<tr>
<td>LSH</td>
<td>lexical shift hypothesis</td>
</tr>
<tr>
<td>M350</td>
<td>an MEG component</td>
</tr>
<tr>
<td>MSE</td>
<td>mean squared error</td>
</tr>
<tr>
<td>MS</td>
<td>mean squared error</td>
</tr>
<tr>
<td>μV</td>
<td>micro volt</td>
</tr>
<tr>
<td>N400</td>
<td>negative ERP component peaking around 400ms post stimulus</td>
</tr>
<tr>
<td>N500</td>
<td>negative ERP component</td>
</tr>
<tr>
<td>N</td>
<td>number of participants</td>
</tr>
<tr>
<td>NMUH</td>
<td>non monotonic update hypothesis</td>
</tr>
<tr>
<td>obj.</td>
<td>direct object</td>
</tr>
<tr>
<td>OBJ</td>
<td>object verb adverbial subject word order</td>
</tr>
<tr>
<td>p</td>
<td>significance level</td>
</tr>
</tbody>
</table>
Acknowledgments

It is difficult to name all the people who have contributed to the work in this monograph. I list only those directly involved: Janina Radó, Fritz Hamm, Wolfgang Sternfeld, Rolf Ulrich, Ingo Hertrich, Hans Kamp and Peter Hagoort. The Collaborative Research Centers SFB 441-Linguistic Data Structures, and SFB 833-Construction of Meaning, deserve special mention. My colleagues provided a stimulating and very pleasing environment. I also thank the participants of our 2005 and 2009 workshops on semantic processing who inspired much of the work reported here. Also, thanks go to all participants of the experiments who spent hours on how to make sense of semantically deviant sentences.

Fritz Hamm inspired much of the work reported here. Janina Radó helped considerably with editing draft versions and discussing the topics right from the start: she also critically evaluated every aspect of this research and helped me a lot in working out my ideas. The research was founded by the German science Foundation and the Collaborative Research center 441 at the University of Tübingen.

This monograph is dedicated to Katrin and to my family to whom goes all my love.
Human language understanding has a very remarkable property. Seemingly effortlessly, we mentally structure the plot of the incoming linguistic material, by encoding and ordering events and actions in our mental construction of time. This is astonishing because the computation of the temporal profile of a described situation rests upon the integration of tense, grammatical aspect and lexical aspectual information as well as world knowledge and information from preceding discourse. This interplay of information at multiple linguistic and extralinguistic dimensions makes temporal processing an excellent test case for investigating semantic interpretation on a supralexical level. This book is about the cognitive mechanisms involved in this construction process during language perception.

The cognitive notion of time as it is reflected in our linguistic categories is only indirectly connected to physical time in its Newtonian representation as a dimension comprising an infinite number of instants corresponding to the real numbers. In contrast, mental time involves the construction of a cohesive representation of ontological primitives like processes, states and events by means of a whole range of relations like precedence, causality, and so forth. Thus, there is a discrepancy between the external and our individual representations of time. Linguistic theory has a lot to offer concerning the latter. Interestingly, crosslinguistic studies on tense and aspect show that although there is variation between languages in terms of how they encode temporal categories and relations, the temporal concepts underlying these varying systems seem to be universal (Dahl 1985; Smith 1991). Moreover, studies on the ontogenesis of temporality in language have shown that English speaking children learn the progressive as their first inflectional form without making stativity-errors at any stage of acquisition (Brown 1973, p. 324–328). That is, children learning English never say things like I am liking you indicating that from the very beginning they cognitively make a distinction between state predicates and non-statives. These findings suggest that the encoding of events in language directly reflect fundamental ontological distinctions between event types. Psychological studies, on the other hand, have shown that event perception can be conceived of as a fundamental cognitive capacity analogous to object perception (Zacks & Tversky 2001).

Although there is a rich theoretical literature on temporality in language, there are only very few psycholinguistic studies on the processing of eventualities. This is
The Processing of Events

astonishing, since the cognitive grounding of temporality in language makes this area particularly interesting for psycholinguistic investigation.

In this monograph, a series of experiments are presented that investigate the mental construction of situation models during language comprehension by looking at processes of so-called aspectual coercion. Aspectual coercion is a reinterpretation of a given situation type into a situation of a different type. The book shows that aspectual coercion affects psycholinguistic measures, for instance reading time, and can thus be used as a real time marker of aspectual semantic processing. The results from the experiments will be used to evaluate a model of aspectual processing. Moreover, I will use the semantic theory of Hamm & van Lambalgen (2005) to model the underlying representations and algorithms employed in aspectual processing.

Chapter 2 provides an introduction to work on events in cognitive psychology and in linguistics. Chapter 3 presents a general model of aspectual processing and states the hypotheses that will be tested in the subsequent chapters. Furthermore, it shows semantic derivations of the relevant constructions and presents an experiment that demonstrates difficulty in aspectual coercion. Chapter 4 describes two reading time experiments that address the question how the processor determines the situation type when it has to deal with ambiguous information. Chapter 5 is about three different kinds of aspectual reinterpretation which will be compared in three reading time studies. Chapter 6 presents two reading time studies and a production experiment that address the question how much of a sentence is needed to compute an event representation. Chapter 7 contains an EEG study on aspectual coercion which discusses the brain correlates of aspectual processing. Chapter 8 summarizes the results and incorporates them into the general model of aspectual processing. Finally, Chapter 9 discusses the findings both with respect to the existing semantic accounts of situation aspect and with respect to the psycholinguistic literature on coercion.

CHAPTER 2

Events in cognitive psychology and linguistics

“All events seem entirely loose and separate. One event follows another; but we never can observe any tie between them. They seem conjoined but never connected. And as we can have no idea of anything, which never appears to our outward sense or inward sentiment, the necessary conclusion seems to be, that we have no idea of connexion or power at all, and that these words are absolutely without any meaning, when employed either in philosophical reasoning, or in private life.”

(Hume 1748), reprinted in Hume (2002, p. 763)

When we read or hear a felicitous piece of discourse we have the intuition that it consists of discrete events which are smoothly connected into a coherent whole. Given the quote from Hume (2002) this seems rather mysterious. How are events glued together? How do we construct the temporal profile of a discourse? How are the pieces brought together in language comprehension? I hope that some of these questions will find an answer in this work.

In this chapter I will look at events from two different perspectives: from a psychological perspective, that is, how we cognitively represent events during scene perception, in memory and in event reporting, and from a linguistic perspective, that is, how different kinds of events are encoded in language. Bringing these two perspectives together will show that many factors considered as central to the cognition of events are reflected in the way language encodes different kinds of events. We will see that the common ground amongst psychologists and linguists includes concepts also relevant for planning in artificial intelligence such as plan-goal structure and hierarchies of events.

2.1 Events from a psychological perspective

Consider the following situation. Hans is sitting at the beach as he spots a ship in distress. Dense smoke is rising from the ship. Five minutes later, the ship has disappeared under water.

If a German speaker wants to describe this situation, he can choose (1a), but the semantically odd (1b) is no option, although in the example watching and sinking coincide.
The Processing of Events

Stein (1997) show that the notion of planning is fundamental for integrating events. Among others, the studies by Trabasso & Nickels (1992) and Trabasso & respect to how humans think and speak about events comes from developmental objects and events can be identified and categorized using hierarchical relations. This location in space and is therefore spatially bounded. Continuing the analogy, both of buttering a toast, for instance, happens at a particular time, but also in a particular event. Analogously, events have boundaries in time, but are also bounded in space. An event active in the spatial domain an event is an ontological primitive in the spatio-temporal domain. Objects are recognized by shape, color etc. and have boundaries in space. Similarly, actions into smaller units. They used an experimental technique which was originally developed by Newton (1973).

The results cannot be explained by a purely perceptual strategy because in their first experiment Zacks et al. (2001) also found an effect of familiarity: the more familiar the action (e.g. washing dishes vs. putting together a saxophone) the higher the alignment of coarse and fine boundaries. Thus, top-down processing helps to identify the internal structure of a complex event.

Another manipulation was included in Zacks et al. (2001)’s first experiment. One group of subjects was instructed to name the event unit they were perceiving while another group was doing the task without naming. Although segmenting and verbalizing at the same time is a dual-task which should be more difficult than non-verbal segmentation, the authors found that the verbalizing group of subjects showed even more alignment than the silent group. This indicates that describing an event helps in accessing its structure. The descriptions were recorded on tape and analyzed with respect to a number of linguistic criteria. 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. Boundary units are the fine unit boundaries that also correspond to coarse unit boundaries. Under the hypothesis that linguistic description of events is also subject to the hierarchical bias effect some correlation in description is expected between the boundary units on the one hand, and the coarse units on the

2.1.1 Experiments on event perception and reporting

Zacks & Tversky (2001) and Zacks et al. (2001) investigate the hierarchical structure of events and subevents which form so-called partonomies. Paying the bill, for instance, is a part (subevent) of a visiting the restaurant event. In a series of experiments they tested whether perceivers make use of partonomic hierarchies to segment a stream of actions into smaller units. They used an experimental technique which was originally developed by Newton (1973).

Participants segmented ongoing activity (like washing dishes) while watching it on film by pressing a key to mark “natural and 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.

The results cannot be explained by a purely perceptual strategy because in their first experiment Zacks et al. (2001) also found an effect of familiarity: the more familiar the action (e.g. washing dishes vs. putting together a saxophone) the higher the alignment of coarse and fine boundaries. Thus, top-down processing helps to identify the internal structure of a complex event.

In Section 2.1.2 we will look at another line of research that is interesting with respect to how humans think and speak about events comes from developmental psychology. Among others, the studies by Trabasso & Nickels (1992) and Trabasso & Stein (1997) show that the notion of planning is fundamental for integrating events into a coherent whole both in narrating stories and in remembering the subevents of a story.
other. And this was indeed the case. For verbs, the authors found that boundary units are described by more general verbs (e.g. *do*) than internal units (e.g. *polish*) and are at an intermediate level between the coarse and the internal units. This provides evidence that even when describing events at a lower level of description, a coarser temporal grain is also being accessed.

Further, there were some common linguistic features between the two classes of unit description. Both used predicates referring to goal-related events rather than providing simple, physically bounded descriptions. For instance, although the participants in the fine unit group could have gone down as far as raises his arm even when providing internal boundary descriptions they stayed at the level of picks up the pot. This suggests that the stream of actions was broken down in units that correspond to intentional acts as they appear in plans without paying attention to their realization at a motoric level. The different levels of granularity are described by Zacks & Tversky (2001, p. 7) as follows:

The smallest psychologically reified events, on the order of a few seconds, may be defined primarily in terms of simple physical changes. For example, think of a person grasping another’s hand, the hands going up, going down, releasing. Longer events, lasting from about 10s to 30s, may be defined in relation to some straightforward intentional act: the events described above, on the time scale indicated, form a handshake. From a few minutes to a few hours, events seem to be characterized by plots (i.e. the goals and plans by their participants) or by socially conventional form of activity. Perhaps the handshake was part of signing a treaty. On time scales that are long enough, it may be that events are characterized thematically. In this example, perhaps the treaty signing was part of an event called a “peace process”. In general, it seems that as the time scale increases, events become less physically characterized and more defined by goals, plans, intentions and traits of their participants.

So far, we have seen how top-down processes do influence event perception. But what are the bottom-up routines employed in event perception? Once again the analogy between objects and events comes into play. An important cue in understanding object recognition is the sudden change in the intensity gradient which marks the boundary between the object and its environment (Marr 1982). Similarly, sudden changes (e.g. in speed or direction) may be important for event perception. I will briefly describe a neuroimaging study that investigated the perceptual basis of event perception.

**Perception systems used in event perception**

Zacks et al. (2006) investigated how movement features are related to the neural processing of events by performing functional magnetic resonance imaging (fMRI) while participants viewed simple animations of two moving objects. After the imaging session, participants watched the animations again and segmented them into fine and coarse meaningful event units in a task similar to the previous studies. Although the objects moved on random trajectories with random speed, participants were able to consistently segment the ongoing activity into discrete event units. Analysis of the movement patterns showed that event boundaries were chosen contingent on speed and direction of motion.

The fMRI data revealed that activation of two specialized brain regions covaried with the placement of event boundaries. The first of these regions, the human MT complex (MT+), is selectively activated by motion. The second region, the posterior superior temporal sulcus (pSTS), indicates the processing of biological motion. Both MT+ and pSTS were selectively activated at those moments in time that participants later marked as coarse-grained perceptual event boundaries. These regions also showed transient changes in activity associated with fine-grained event boundaries although the effect was less robust. Taken together, the study establishes that brain regions involved in processing motion in general and biological motion in particular respond selectively at the boundaries between events. The analysis of motion features by the brain may thus provide a causal pathway of how events are perceived.

### 2.1.2 The importance of planning in narrating and remembering events

The connection between integrating events in a coherent whole and being goal-oriented was investigated by Trabasso & Nickels (1992) and Trabasso & Stein (1997). Based on developmental data they come to the conclusion that causality and planning provide the medium through which isolated events become glued together into a larger whole.

Trabasso & Nickels (1992) present the findings of an experiment in which children and adults narrated a sequence of 24 scenes in a picture storybook called *Frog, Where Are You?*, in which a boy searches for and finally finds his pet frog which has escaped from its jar. The authors analyzed the linguistic devices that were used to narrate the story as a function of age.

The youngest group were three-year old children. They tended to tell the story in a tenseless fashion, describing single objects and actions without relating them to each other or to the boy’s ultimate goal. At age four, children began to use temporal expressions like *then* to sequence the actions temporally and sometimes chose actions that are relevant to a goal. But when they did mark actions purposefully, these were local and dependent on the picture. At the age of five years a different picture emerges. The narrations indicate explicit awareness that a particular action is instrumental towards a goal. At age nine, goal-attempt-outcome descriptions become more frequent than non-goal-directed descriptions and this tendency becomes a little stronger in adult narrations.

The experiment provides evidence for a gradual development in the ability to structure the narrative as the execution of a plan towards the goal of finding the frog.
This seems to be completely impossible for the three year olds who appear to be entirely glued to the current state. Only at the age of five, can the child move from the current action to mentioning the goal and back again. But why does this ability develop so late? The following quotation from Trabasso & Stein (1994, p. 327) may provide an explanation by describing the cognitive demands of the task:

Inferring goals and plans involves considerable social and personal knowledge and places heavy demands on a narrator's working memory. The child who narrates events needs to attend to and maintain the current event in working memory; to activate and retrieve prior knowledge relevant to events, either in general or from earlier parts of the story, in order to interpret and explain the current event; and to integrate these interpretations into a context within a plan, all within the limitations of knowledge and working memory. In effect, over time the child is engaged in dynamic thinking, actively constructing and evaluating models and hypotheses about what is occurring. In so doing, the child creates a changing mental model that results in a long-term memory representation of what has occurred.

The second piece of evidence for the importance of planning in understanding complex events comes from an experiment on event memory by Trabasso & Stein (1997). Again, picture stories were used (consisting of only six pictures this time) that had to be narrated by children. After a child had narrated the story s/he was asked to recall the pictures of the story from memory. The trials were divided in two categories: trials in which the child had produced a modal description, that is a description keeping with the goal structure of the story, and non-modal trials where the description wasn't connected to the ultimate goal. The retention rates showed an interesting difference between modal and non-modal trials. While the non-modal trials showed a clear primacy and recency effect – that is the first and the last pictures could be remembered more often than the intermediate pictures, a well known effect in research on memory – in the modal trials there was no primacy and no recency effect. This shows that integrating events into a goal structure leads to sustained activation of the complete event sequence which is now junked together into a complete whole.

2.1.3 A brief summary

The above mentioned studies show that events are cognitively real. They further provide evidence that our mental representations of events are organized in partonomic hierarchical structures. In describing events, we make use of intentions and goals of the event participants and use language that reflects a close connection to planning. This directly leads us to the linguistic realm where event descriptions have been extensively studied. In the course of the discussion I will come back to the implications of the studies mentioned above and show that semantic theory contributes a formal theoretical framework which can deal with all the psychologically relevant aspects of event representations.

2.2 Events in linguistics

In linguistics and philosophy there is a long-standing tradition dating back at least to Davidson (1967) that takes events seriously and assumes them to be a fundamental semantic concept. In event-semantics they are thought to be ontologically primitive and form the basic unit in the meaning of action sentences like (2a) and (2b).

(2) a. John sneezed.
   b. John sneezed loudly.

Davidson argues that without introducing events into the logical form of the sentences it is rather mysterious why (2b) entails (2a). To represent them in classical first order predicate logic we need two different predicates, a one-place predicate (3a) and a two-place predicate (3b).

(3) a. sneeze, (john)
   b. sneeze, (john, loud)

Assuming that this representation is correct how can the entailment patterns be determined? This is not a simple issue as becomes clear by comparing John sneezed five minutes ago and John sneezed three minutes ago where neither sentence is implied by the other. The logical form must explain these facts. One possibility is to formulate a meaning postulate like (4) to enforce the entailment.

(4) ∀x.∀y. sneeze(x, y) → sneeze, (x)

But the meaning postulate (4) is just an ad hoc solution. As Davidson further argues adding adverbials to an action sentence makes the entailment patterns increasingly complex. In principle, an infinite number of meaning postulates is needed for each verb to get things right. This is illustrated by Davidson's famous Example (5).

(5) a. Jones buttered the toast in the bathroom with a knife at midnight.
   b. Jones buttered the toast in the bathroom with a knife.
   c. Jones buttered the toast in the bathroom.
   d. Jones buttered the toast.

1. Reichenbach (1947) was even earlier in proposing an event semantics for adverbial modification.
A solution to this dilemma is to represent even (2a) in a logically more complex way than sneeze\( (\text{john}) \). According to this solution, sentences that talk about actions or events make explicit reference to an event argument. The adverbial phrases are then added by conjunction. (6a) and (6b) show the logical form of (2a) and (2b).

\[
\begin{align*}
\text{(6) a. } & \exists e. \text{sneeze}(e, \text{john}) \\
\text{b. } & \exists e. \text{sneeze}(e, \text{john}) \land \text{loud}(e)
\end{align*}
\]

It is clear that the representation in (2a) follows from (2b). On the other hand, no entailments hold between John sneezed five minutes ago and John sneezed three minutes ago because the latter two sentences talk about different events. Davidson’s work shows that from a philosophical/linguistic perspective it makes good sense to include event types in the ontological repertoire of Semantics.

Should all action sentences be treated ontologically in a uniform way? More concretely, should John sneezed and John ran be represented alike? In an influential paper Bach (1986) proposed an algebra of events which draws a categorical distinction between two sorts of eventualities – events and processes. Each of them forms an algebra and these two algebras are connected by a homomorphism from the algebra of events to the algebra of processes. The algebras are different with respect to the sets and the operations. The same distinction has been made by Link (1983) in the nominal domain between count and mass nouns, for instance, the ring versus the gold in the ring. For simplicity, I will first present an example from the nominal domain and then move on to events and processes.

Suppose my ring is hundred percent gold. Then, my ring and the gold in my ring are physically identical. Nevertheless, the sentences (7a) and (7b) can both be true although the predicates are mutually incompatible.

\[
\begin{align*}
\text{(7) a. } & \text{My ring isn’t very old.} \\
\text{b. } & \text{The gold in my ring is very old.}
\end{align*}
\]

This shows that the two entities my ring and the gold in my ring aren’t the same. While the former refers to a thing, the latter refers to an amount of stuff. For things all set-theoretic operations are allowed: union (e.g. my ring and your ring form the set of our rings), intersection (e.g. the rings in this drawer) and complement (e.g. All my things except for the ring). Also, there is a partial ordering \( \preceq_{\text{thing}} \) defined for things, that is, John is a part of the plural entity John and Mary. More technically, the domain of things is a complete Boolean algebra. Stuff, on the other hand, is a structured set of quantities of objects with join operation and an ordering relation \( \preceq_{\text{stuff}} \). One can think of this as what one gets by taking every quantity of stuff and further dividing it up into all its subquantities. Five liters of water, for example, can be divided up into two and three liters. Technically speaking, stuff forms a join semilattice, that is an algebra rather different from the algebra of things. But objects and their constitutive quantities of stuff are related by a homomorphism \( h \). The homomorphism guarantees that whenever two things a and b stand in the order relation \( a \preceq_{\text{thing}} b \) then their constitutive quantities of stuff \( h(a) \) and \( h(b) \) also stand in the relation \( h(a) \preceq_{\text{stuff}} h(b) \).

A similar relationship can be observed between events and processes. While only processes co-occur with expressions denoting amounts (like a lot), only events co-occur with quantifiers denoting numbers (like forty times). In this regard, they behave exactly like mass nouns and count nouns. Consider the analogy between (8) and (9).

\[
\begin{align*}
\text{(8) a. } & \text{John ran a lot last year.} \\
\text{b. } & \text{John ran ten miles a lot last year.} \\
\text{c. } & \text{John ran forty times last year.} \\
\text{d. } & \text{John ran ten miles forty times last year.}
\end{align*}
\]

\[
\begin{align*}
\text{(9) a. } & \text{A lot of gold was found in the mine.} \\
\text{b. } & \text{A lot of ring is made out of gold.} \\
\text{c. } & \text{At least forty of these golds were found in the mine.} \\
\text{d. } & \text{At least forty of these rings are made out of gold.}
\end{align*}
\]

Bach (1986) argues that analogously to nouns eventualities consist of two algebras of different types, an atomic Boolean algebra of events and a join semilattice of processes. Just like in the nominal system the two algebras are connected by a homomorphism. Events and processes have to be kept ontologically separate: events are countable, processes are measurable in time. In the following I will follow Hamm & van Lambalgen (2005) and use the terms event and fluent to distinguish between entities of these two sets.

\[
2.2.1 \text{ A calculus of events (Hamm \& van Lambalgen (2005))}
\]

The experimental results from cognitive psychology substantiate the claim that planning is fundamental for the mental construction of events. This idea has been carried over to natural language semantics by Hamm \& van Lambalgen (2005) who proposed the Event Calculus (EC), a planning formalism that was originally developed for path planning in robotics. In the following sections I will introduce important aspects of the formalism piece by piece. A much more complete picture can be found in Hamm \& van Lambalgen (2005) from which much is borrowed here.
The purpose of planning is to find a sequence of actions to achieve some objective when starting from some given situation. Planning is the reasoning side of acting and involves reasoning about events – both actions of an agent and events in the environment – and about properties of the agent and the environment. The latter may change as a consequence of the events. Planning is a form of non-monotonic or default reasoning. A plan can never be proven to be correct in the sense of classical logic, that is in every model of the premises, but can always fail due to unforeseen obstacles. In fact, it may be impossible to enumerate all the things that can go wrong. As far as the formalism is concerned, what is needed instead of a provably correct plan is a plan that works to the best of one's knowledge. EC uses a non-monotonic logic, namely logic programming, employing the closed world assumption which states that everything one has no information about can be assumed to be false. Adding further information may thus enforce a new computation that is in conflict with the initial plan.

Let's look at a simple example. Suppose an agent wants to get from room 1 into room 2 and knows there is a door connecting the two rooms which is shut. The obvious plan would be to open the door and go through into room 2. However, if the door turns out to be locked the original plan fails. Adding the information about the locked door to the database makes it necessary to recompute the plan and add further actions to it, e.g. to get the key first and unlock the door. Hence, a successful planning algorithm must be flexible and allow for recomputations.

What does path planning have to do with language interpretation? A listener has an initial discourse model and upon hearing a new sentence he must integrate the new sentence into. He has to update that model with the actions and the participants stated in the sentence and has to keep track of the consequences of these actions. Understanding a discourse like the frog story can thus be viewed as coming up with a plan to achieve the ultimate goal (e.g. how can the boy get hold of his frog again?). Similarly to the robotics example, there may be obstacles mentioned later in the discourse that call for a recomputation of the discourse model. Such non-monotonic updates play an important role in EC.

**Basic ontology**

EC is a many sorted first order logic. It includes the following sorts:

- objects and individuals, like John and the chair;
- real numbers, to represent time and variable quantities;
- fluents, symbolized as $f, f', f''$, ..., which are time-dependent properties like states (c.f. being on top) and activities (c.f. walking), but also variable quantities, such as spatial position (= position($x_1, x_2, x_3, t$)) or state of completion of a house (= house($x, t$));
- event types, symbolized as $e, e', e''$, ..., whose tokens typically mark the beginning and end of fluents (e.g. sneeze(John)).

**Primitive predicates**

EC has a set of primitive predicates for handling two types of causality – instantaneous and continuous change. The former can be exemplified by the classical example of two balls colliding. The paradigmatic example for the latter is the acceleration of an object in a gravitational field. Other examples are pushing a cart and filling the bucket. Let's have a closer look at both types of change.

Imagine a situation of ball 1 moving while ball 2 is lying motionless on the table. At time 5 ball 1 hits ball 2, terminating the movement of ball 1 and causing ball 2 to move. To characterize this situation we need a description of the initial state, a statement about an event type collision and the consequences of this event. Let $f$ be a variable over fluents and $e$ over events.

- Initially($f$) expresses that fluent $f$ holds at the beginning of the situation.
- Happens($e, t$) means that event type $e$ has a token at $t$.
- Initiates($e, f, t$) says that the causal effect of $e$ at time $t$ is the fluent $f$.
- Terminates($e, f, t$) expresses that the causal effect of $e$ at time $t$ is the negation of fluent $f$.

The exact meanings of the primitive predicates are defined by the axiom system of EC which will be described in the next section. Using the basic predicates (10) formalizes the event of the two balls colliding.

(10) Initially(move(ball1))
     Initially(~move(ball2))
     Happens(collision, 5)
     Initiates(collision, move(ball2), t)
     Terminates(collision, move(ball1), t)

The second notion of causality is continuous change that is due to a force. I will use the example of a bucket that contains one liter of water at time 0 and after a tap is turned on at time 1 becomes filled with one liter per time unit. To formalize the changing content of water two special predicates are required.

- Trajectory($f_1$, $t$, $f_2$, d) where $f_1$ is a force and $f_2$ is a variable quantity which may change under the influence of the force.
- Releases($e, f, t$) is necessary to bring the two notions of change together. The predicate says that $e$ at time $t$ releases $f$ from the common sense law of inertia which ensures that once the initiating event happened the fluent is free to vary continuously. By contrast, purely instantaneous change "freezes" the world in its consequent state. The common sense law of inertia is one solution for the notorious frame problem known from AI (McCarthy & Hayes 1969): Using mathematical logic, how is it possible to write formulæ that describe the effects of actions without having to write a large number of accompanying formulæ that describe the mundane, obvious non-effects of those actions?
The formalization of the “filling the bucket with water” situation in (11) is straightforward. We need an event \(\text{turn on that releases the fluent water}(x)\) representing a variable quantity. The fluent \(\text{water}(x)\) is driven by a fluent \(\text{filling that acts as force. If filling}\) is true between times \(t\) and \(t + d\), then at time \(t + d\) the bucket is filled to degree \(x + d\).

\[
(11) \quad \text{Initially(water}(1))
\]
\[
\text{Happens(turn on, 1)}
\]
\[
\text{Releases(turn on, water}(x), t)
\]
\[
\text{Trajectory(filling, t, water}(x + d), d)
\]

Apart from primitive predicates that deal with the representation of change, there are two other basic predicates. \(\text{Clipped}(t, f, t')\) means that there is a “\(t\)-relevant” event occurring between \(t_1\) and \(t_2\) that is a terminating or releasing event. In our first example \(\text{Clipped}(4, \text{move(ball2)}, 6)\) is true because the terminating event \(\text{collision}\) happened in the time interval between time 4 and time 6. \(\text{Clipped}\) will play an important role in stating the axioms in the next section. The second predicate, \(\text{HoldsAt}(f, t)\), serves as a truth predicate and says that \(f\) holds at time \(t\).

**Axioms of the Event Calculus**

EC is an axiatomized theory. The axioms provide a general theory of temporality that embodies the principle of inertia. Below is the list of axioms. All variables are assumed to be universally quantified and the axioms are stated in the form of logic programming (PROLOG) clauses. Appendix A contains a brief introduction to (propositional) logic programming and shows how logic programs can be translated into (recurrent) neural nets.

**Axiom 1** \(\text{Initially}(f) \rightarrow \text{HoldsAt}(f, 0)\)

**Axiom 2** \(\text{Happens}(e, s) \land t < s < t' \land (\text{Terminates}(e, f, s) \lor \text{Releases}(e, f, s)) \rightarrow \text{Clipped}(t, f, t')\)

**Axiom 3** \(\text{Happens}(e, t) \land \text{Initializes}(e, f, t) \land t < t' \land \neg \text{Clipped}(t, f, t') \rightarrow \text{HoldsAt}(f, t')\)

**Axiom 4** \(\text{Happens}(e, t) \land \text{Initializes}(e, f_1, t) \land t < t' \land t' = t + d \land \text{Trajectory}(f_1, t, f_2, d) \land \neg \text{Clipped}(t, f_2, t') \rightarrow \text{HoldsAt}(f_2, t')\)

Axiom 1 defines the predicate \(\text{Initially}(f)\). The predicate is true if the fluent is true at time 0, that is at the beginning of a given situation. Axiom 2 defines the predicate \(\text{Clipped}(t, f, t')\) which is true if \(f\) is terminated or released by an event \(e\) which happens in the time interval between \(t\) and \(t'\). Thus \(\text{Clipped}\) covers both instantaneous and continuous change. Axiom 3 embodies the principle of inertia. If an event initiates a fluent \(f\) at time \(t\) and no \(f\)-relevant event happens later on, the fluent is supposed to hold at all times \(t'\) after \(t\). Axiom 4 defines the \(\text{Trajectory}\) predicate. To see what it says, let’s consider the example of filling the bucket once more where \(f_1\) is instantiated by \(\text{filling}\) and \(f_2\) by \(\text{water}(x)\). Should \(\text{filling}\) be true during the whole interval from \(t\) until \(t'\) then for a certain \(x\), \(\text{water}(x)\) will be true at \(t'\). The value of \(x\) will be determined by the law of the process under consideration (e.g. filling in one liter per time unit). The axioms provide a general theory of what can change and what stays constant. They hold for every situation no matter what it may be. For concrete situations, we also need the specific causal relationships. This kind of information is specified in so called scenarios. Let’s move on to these.

**Scenarios**

Scenarios are micro-theories which express lexical meaning. They state the events happening in a concrete situation, the fluents holding at a given time and the causal relationships between the eventualities. When we looked at representations for instantaneous and continuous change in (10) and (11) we have already worked with two scenarios. In general, a scenario is a conjunction of statements that have the form of facts or rules in logic programming, for instance \(\text{S}(t) \rightarrow \text{Trajectory}(f_1, t, f_2, d)\) where \(\text{S}(t)\) is a state which roughly corresponds to a positive or negative \(\text{HoldsAt}\) predicate. (12) is the complete scenario for \(\text{filling the bucket}\). The last statement of (12) expresses the rule that if the bucket contains a quantity of \(x\) water at time \(t\) it will contain \(x + d\) water after \(t + d\) time.

\[
(12) \quad \text{Initially(water}(1))
\]
\[
\text{Happens(turn on, 1)}
\]
\[
\text{Releases(turn on, water}(x), t)
\]
\[
\text{HoldsAt(water}(x), t) \rightarrow \text{Trajectory}(\text{filling}, t, \text{water}(x + d), d)
\]

Scenarios differ depending on the situation type or Aktionsart expressed by the sentence. While some contain statements about instantaneous and/or continuous change others describe stative situations that do not involve any change at all. In the following

\(\text{is a bird then } x \text{ flies}\) has a non-monotonic interpretation because an extension of the theory does not necessarily preserve truth. For instance, if the database contains the fact that \(\text{Tweety}\) is a bird, PROLOG will answer “yes” to the query \(\text{flies(Tweety)}\). However, adding the fact that \(\text{Tweety}\) is a penguin together with the rule \(\text{flies}(X) \rightarrow \text{penguin}(X), \text{!}, \text{fail}\) stating that penguins do not fly, will change the truth value from true to false and the query \(\text{flies(Tweety)}\) will be answered “no”.

---

3. EC includes one more axiom which guarantees that all fluents \(f\) that hold for the first time also hold at later times until a \(f\)-relevant event happens. For simplicity, this axiom was left out here.

4. The \(\rightarrow\) should not be taken to be the material conditional of classical logic, but correspond to rules in PROLOG, a programming language that can be used for planning (for an introduction see Bratko (2001)). A PROLOG rule like \(\text{flies}(X) \rightarrow \text{bird}(X)\) meaning for all \(x\), if \(x\) is a bird then \(x\) flies has a non-monotonic interpretation because an extension of the theory does not necessarily preserve truth. For instance, if the database contains the fact that \(\text{Tweety}\) is a bird, PROLOG will answer “yes” to the query \(\text{flies(Tweety)}\). However, adding the fact that \(\text{Tweety}\) is a penguin together with the rule \(\text{flies}(X) \rightarrow \text{penguin}(X), \text{!}, \text{fail}\) stating that penguins do not fly, will change the truth value from true to false and the query \(\text{flies(Tweety)}\) will be answered “no”.
section I will present a typology of situation types in language and will provide typical scenarios for each of them.

### 2.2.2 The representation of Aktionsart

Although the study of lexical aspect has a long history starting already with Aristotle, it was Vendler (1957, 1967)'s four-way classification into states, activities, accomplishments and achievements that proved to be most influential in linguistics (for an extended discussion see Dowty (1979) or Rothstein (2004)). According to this classification, states like *be happy* can be characterized as non-dynamic situations while activities like *run* involve open-ended processes, and accomplishments like *build a house* are telic processes with a natural end-point. The last aspectual class in the Vendlerian taxonomy, achievements like *die* or *reach the top*, involve instantaneous change. Building upon this classification Dowty (1979) formulated his seminal semantic analysis and discussed the logical entailments of these classes, their interaction with temporal modifiers and with tense. Smith (1991) introduced a fifth aspectual class, semelfactives or points like *flash* which express an event, but do not involve any change of state. The world is thought to stay the same after a flashing event occurred, whereas in the case of an achievement like *die* the event can change the state of affairs rather drastically. This classification into five Aktionsarten should be seen as a non-linguistic conceptual typology, although it is often called lexical aspect. This is because in languages like English or German the categories cannot be systematically mapped to lexical items. Instead, lexical items can almost always be coerced into a different Aktionsart.

There have been numerous attempts to describe the aspectual classification in a feature-based manner. For instance, according to Rothstein (2004) the four-way Vendlerian expression intuitions that there are two properties which are crucial for categorizing eventualities or event types. These two properties are whether an event type is telic and whether we can analyze it as progressing or developing. While telicity is well acknowledged in other feature-based classifications as well (e.g. de Swart (1998)), the second property is not so commonly assumed. The first property, [+ telic] groups states and activities together on the one hand and accomplishments and achievements on the other. The endpoint of a telic predicate is often called its culmination, which consists in a punctual event. By contrast, an atelic predicate lacks a culmination and can in principle last infinitely long. For instance, *die* or *reach the top* as its natural endpoint, but *run* has no culmination. Rothstein (2004, p. 8) modifying the account of Krifka (1998), gives the following characterization of telicity: if an event e is in the denotation of a predicate X, then all parts of e which are also in the denotation of X have the same end point. The classical test for telicity is adverbial modification by *for x time* vs. *in x time*. Telic predicates can easily combine with in adverbials, but are generally bad with *for* adverbials while atelic predicates show the opposite pattern. This is illustrated in (13).

(13) a. John knew her for twenty years/*in twenty years. (= state)
   b. John ran for twenty minutes/*in twenty minutes. (= activity)
   c. John built a house *for one year/in one year. (= accomplishment)
   d. John reached the top *for two hours/in two hours. (= achievement)

The second property, [+ stages] allows for an orthogonal grouping of aspectual classes: While both accomplishments and activities involve continuous processes, states and achievements are essentially non-dynamic. The prototypical test for this property is the progressive test. Only predicates that involve stages can be put in the progressive as is demonstrated by the sentences in (14).

(14) a. *John is loving his mother. (= state)
   b. John is jogging in the park. (= activity)
   c. John is building a house. (= accomplishment)
   d. ?John is finding his keys. (= achievement)

[± telic] and [± stages] allow a systematic treatment of the four classical aspectual classes. However, the existence of semelfactive predicates as a fifth aspectual class disrupts this classification. The main motivation for introducing features is to explain why we have exactly the classes we have. Using n binary features one would expect to have realized all 2^n combinatory possible combinations of feature values. Trying to fit semelfactives into the feature-based classification proves to be difficult because semelfactives are neither telic nor atelic as shown in (15).

(15) a. *The balloon popped for five minutes.
   b. *The balloon popped in five minutes.

Instead of adopting a purely feature-based representation of aspect, I will follow Moens & Steedman (1988), Pulman (1997) and Hamm & van Lambalgen (2005) in choosing the event nucleus as the representation of the aspectual classes. These I will treat as ontological primitives.

[The event nucleus is] an association of a goal event, or culmination, with a preparatory process by which it is accomplished, and a consequent state, which ensues. (Moens & Steedman 1988, p. 15)

Features will only play a role at the lexical level to determine Aktionsart in a compositionally transparent way. The event nucleus is a complex semantic representation encoding temporal profiles of the aspectual classes and consists of processes, states, events and time-dependent properties. It relates Aktionsart directly to cognition, namely in being a minimal instantiation of a plan: it connects an initial state to the desired consequent state via a course of actions that leads to a change in the intended
The Processing of Events

Chapter 2. Events in cognitive psychology and linguistics

The incremental theme in the sense of Dowty (1979) is not represented in the classical event nucleus by Moens & Steedman (1988), but has been added by Hamm & van Lambalgen (2005). Consider an accomplishment like build a house. The building process is acting as a force on a continuously changing object, namely the house which gradually comes into existence. Hamm & van Lambalgen (2005, p. 84) define the general concept of an eventuality like follows:

**Definition 5** An eventuality is a structure \( \langle f_1, f_2, e, f_3 \rangle \), where

1. \( f_1 \) is a fluent which represents an activity, something which exerts a force
2. \( f_2 \) is a parameterized fluent, representing a parameterized 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) or ‘f’ (first, second and fourth argument).

One advantage of the representation format of the event nucleus has to do with differences in acceptability for sentences like (13a–d) and 14a–d). While some of the sentences are clearly unacceptable, others are quite acceptable, at least in combination with the right context. This is due to the possibility of type shifting which will be discussed in a later section. The binary nature of a feature-based system only allows for dichotomous differentiations, while the cognitively grounded setup of the event nucleus can explain how aspectual mismatch can be resolved. Let's look at the event nuclei of the five aspectual classes in turn.

**States (eventualities of type \( \langle -, -, -, f_3 \rangle \))**

States are typically expressed by stative verbs like like, be sad, resemble, know etc. They are conceived of as homogeneous and durative situations. Their internal temporal nature is totally undifferentiated, that is, although extended in time a state doesn't involve any dynamics and every part of it is itself a state of the same type. States are causally inert, they cannot exert a force and therefore cannot appear as first argument \( f_1 \) of an eventuality. In a now standard description, Comrie (1976, p. 49) characterizes states as follows:

With a state, unless something happens to change that state, it will continue: this applies equally to standing and to knowing. With a dynamic situation, on the other hand, the situation will only continue if it is continually subject to a new input of energy: this applies equally to running and to emitting a pure tone, since if John stops putting effort into running, he will come to a stop, and if the oscilloscope is cut off from its source of power it will no longer emit sound. To remain in a state requires no effort, whereas to remain in a dynamic situation requires effort, whether from the inside (in which case we have an agentive interpretation, e.g. john is running), or from the outside (in which case we have a nonagentive interpretation, e.g. the oscilloscope is emitting a pure tone).

In EC, a state is formally represented by a positive or negative \( \text{HoldsAt} \) predicate. The primitive predicate \( \text{HoldsAt}(f, t) \) serves as truth predicate and states that the fluent \( f \) is true at time \( t \). (16a) provides an example of a state and its translation (16b) into EC.

\[
(16)\ a. \ \text{Mary is}(\text{n't}) \ \text{sad}.
\]
\[
(16)\ b. \ (\neg)\text{HoldsAt}(\text{sad(mary)}, t)
\]

There are a number of linguistic tests that can be used to identify states, summarized in Table 2.1. The most important is that when used in English in the present, states seem to be incongruous with the progressive. This is shown by the contrast in (17).

\[
(17)\ a. \ \text{John is sick}.
\]
\[
(17)\ b. \ \neg \text{John is being sick}.
\]
\[
(17)\ c. \ \text{John knows the answer}.
\]
\[
(17)\ d. \ \neg \text{John is knowing the answer}.
\]

It's important to note that the apparent ungrammaticality only holds for the default interpretation of the stative predicates without further context. Consider (18) (which I googled from the internet).

\[
(18)\ \text{One of the biggest thrills of my young life was knowing the answer to a riddle one day.}
\]

(18) describes a quiz context where the speaker had to come up with the correct answer which was perceived as an effortful task. In accordance with observations like this, Hamm & van Lambalgen (2005) hypothesize that the difference between states and dynamic situations lies in the cognitive perspective taken upon the eventuality. If a situation is viewed as an activity the input of energy is profiled, whereas when viewed as a state the result of that input is in the focus. It is thus possible to shift emphasis from the result to the corresponding activity as shown in (18) and (19), our first examples of aspectual coercion. In the latter, love appears in the progressive and the stative predicate is coerced into a dynamic eventuality.

\[
(19)\ \text{John is loving Mary more and more, the more he gets to know her.}
\]
States can be further divided in permanent versus temporary ones. While be tall is an example of the first category, be sick is an instance of the second. Some languages go so far as to morphologically mark the difference between them, as does Kobon, a language spoken in Papua New Guinea which only has a relatively small number of verbs (see Davies (1981) and Hindsill (2007, p. 16–19) for an interesting discussion of the data).

Activities (eventualities of type \( \langle f_1, f_2, \ldots, f_n \rangle \))

Activities are characterized by being dynamic and energy consuming. In contrast to accomplishments, they lack a canonical goal. Thus, without having a certain distance in mind, one cannot say John finished running, but has to say John stopped running. We can distinguish between activities in the strict sense (the boy was trembling) and in the wide sense (the boy was running). Only the latter contain what Dowty (1991) called the "incremental theme"; e.g. the distance the boy has run. In the following I will only focus on activities in the wide sense.

EC represents activities (in the wide sense) by a force \( f_1 \) (= activity in the narrow sense) which leads to a change in the incremental theme \( f_2 \). Change comes about by the trajectory predicate which relates \( f_1 \) to \( f_2 \) via some function \( g \). In (20), \( g \) is a monotone increasing real-valued function relating the pushing activity (= push\((j,c)\)) to the position of the cart (= pos). The value \( a \) represents the initial position of the cart in the scenario (20b). Via releases, the position of the cart is allowed to vary, it is no longer subject to the common law of inertia. The whole scenario is stated as a logic program. Tense and (progressive) aspect are left aside for the moment.

(20) a. John is pushing a cart.
   b. \[
   \begin{align*}
   \text{John push- a cart} \\
   1. \text{Initially(pos}(a)) \\
   2. \text{Initiates} (\text{start}, \text{push}(j,c), t) \\
   3. \text{Releases} (\text{start}, \text{pos}(x), t) \\
   4. \text{HoldsAt(pos}(x), t) \rightarrow \text{Trajectory} (\text{push}(j,c), t, \text{pos}(x + g(d)), d)
   \end{align*}
   \]

There are two standard tests for activities: first, activities are bad with container adverbials (in x time), but good with durative adverbials (for x time). This was already shown in (13b). Second, in contrast to accomplishments, they aren’t subject to the so called imperfective paradox. That is, in the progressive they show different entailment patterns than accomplishments. Compare the following sentences.

5. And thus conceiving running as an accomplishment.

Activities (eventualities of type \( \langle f_1, f_2, \ldots, f_n \rangle \))

Activities are characterized by being dynamic and energy consuming. In contrast to accomplishments, they lack a canonical goal. Thus, without having a certain distance in mind, one cannot say John finished running, but has to say John stopped running. We can distinguish between activities in the strict sense (the boy was trembling) and in the wide sense (the boy was running). Only the latter contain what Dowty (1991) called the "incremental theme"; e.g. the distance the boy has run. In the following I will only focus on activities in the wide sense.

EC represents activities (in the wide sense) by a force \( f_1 \) (= activity in the narrow sense) which leads to a change in the incremental theme \( f_2 \). Change comes about by the trajectory predicate which relates \( f_1 \) to \( f_2 \) via some function \( g \). In (20), \( g \) is a monotone increasing real-valued function relating the pushing activity (= push\((j,c)\)) to the position of the cart (= pos). The value \( a \) represents the initial position of the cart in the scenario (20b). Via releases, the position of the cart is allowed to vary, it is no longer subject to the common law of inertia. The whole scenario is stated as a logic program. Tense and (progressive) aspect are left aside for the moment.

(20) a. John is pushing a cart.
   b. \[
   \begin{align*}
   \text{John push- a cart} \\
   1. \text{Initially(pos}(a)) \\
   2. \text{Initiates} (\text{start}, \text{push}(j,c), t) \\
   3. \text{Releases} (\text{start}, \text{pos}(x), t) \\
   4. \text{HoldsAt(pos}(x), t) \rightarrow \text{Trajectory} (\text{push}(j,c), t, \text{pos}(x + g(d)), d)
   \end{align*}
   \]

There are two standard tests for activities: first, activities are bad with container adverbials (in x time), but good with durative adverbials (for x time). This was already shown in (13b). Second, in contrast to accomplishments, they aren’t subject to the so called imperfective paradox. That is, in the progressive they show different entailment patterns than accomplishments. Compare the following sentences.

5. And thus conceiving running as an accomplishment.

Accomplishments (eventualities of type \( \langle f_1, f_2, f_3, \ldots, f_n \rangle \))

Like activities, accomplishments describe dynamic situations. But in addition they contain a canonical goal which finishes the preparatory process. To determine whether a sentence expresses an activity or an accomplishment much depends on the semantic properties of the arguments. This is illustrated by the contrast in (22a) versus (22b) where the only difference lies in the mass/count status of the direct object. (22a) with the mass noun water expresses an activity while (22b) with the count NP a glass of water is an accomplishment.

(22) a. John is drinking water.
   b. John is drinking a glass of water.

The EC representation of an accomplishment contains an activity as a proper part and adds some statements which involve the goal finish. (23a) is an accomplishment and (23b) its EC representation, again leaving tense and progressive aspect aside.

(23) a. John is building a house
   b. \[
   \begin{align*}
   \text{John build- a house} \\
   1. \text{Initially(house}(a)) \\
   2. \text{Initiates} (\text{start}, \text{build}(j, h), t) \\
   3. \text{Initiates} (\text{finish}, \text{house}(c), t) \\
   4. \text{Terminates} (\text{finish}, \text{build}(j, h), t) \\
   5. \text{HoldsAt(build}(j, h), t) \wedge \text{HoldsAt(house}(c), t) \rightarrow \text{Happens} (\text{finish}, t) \\
   6. \text{Releases} (\text{start}, \text{house}(x), t) \\
   7. \text{HoldsAt(house}(x), t) \rightarrow \text{Trajectory} (\text{build}(j, h), t, \text{house}(x + g(d)), d)
   \end{align*}
   \]

The accomplishment (23a) starts with an activity build\((j, h)\) that causes the partial object house\((x)\) to change. The arguments of build, that is \( j \) and \( h \), serve as abbreviations for John and the house. At some point, the house is completed (= house\((c)\)). This triggers a finish event that terminates the building activity.

6. The complete list of tests can be found in Table 2.1.
The classical test that distinguishes accomplishments and activities is their behavior in connection with container and durative adverbials (see (13c)). Another difference is that accomplishments in the progressive are subject to the imperfective paradox. This was illustrated in (21b). Finally, as the reader may expect, accomplishments can be embedded in sentences with finish, but are semantically deviant with stop. Consider the contrast between (24a) and (24b).

(24) a. John finished building a house.
   b. ?John stopped building a house.

Finally, there is an ambiguity with almost used with accomplishments that doesn't occur with the other Aktionsarten. Take the sentences in (25).

   b. John almost built a house.

The activity sentence (25a) entails that John didn't push the cart, while the accomplishment (25b) can either mean that John almost started the building activity or that he almost finished building the house (for a complete list of tests see Table 2.1).

The distinctions between the aspectual classes aren't always clear cut. In fact, Aktionsarten are more like prototypes and particularly activities and accomplishments are divided by a fuzzy boarder. For instance, the contrast between (24a) and (24b) is gone if we use another predicate instead of run a mile:

(26) a. John finished writing an essay.
   b. John stopped writing an essay.

German has prefix verbs that are close to unambiguous accomplishments. For instance, einen Berg besteigen (to climb a mountain) or ein Haus errichten (to build a house) have a very predominant accomplishment reading. Accomplishments like these will play an important role in the experiments that will be described in the following chapters.

Semelfactives (eventualities of type \(-\_\,-, e, f\_\))

The event types discussed thus far have all been durative. Let's move on to events in a more natural sense of the word, namely eventualities that involve abrupt rather than gradual change. An event in the strict sense is something that happens instantaneously. (27) is an example of a semelfactive, that is, a type of situation only comprising a single event. In the EC representation, the coughing event appears as the argument of the primitive Happens(cough) predicate which states that a coughing event e happens at some instant of time t.

(27) a. John coughed.
   b. Happens(cough(john), t)

Objectively, virtually nothing happens strictly instantaneously. A cough, for example, takes time and can be broken down into phases when viewed in slow motion. The important point here is that linguistic distinctions reflect cognitive dimensions which need not correspond to objective reality.

There are linguistic reasons to assume that semelfactives denote punctual events. If a speaker wants to refer to a cough in a durative manner (e.g. while watching a slow-motion movie about the respiratory system), he will probably use an imperfective form like now he is coughing. There is also typological evidence for the distinction between punctual events and events that make up an interval of time. Russian, for instance, has verb classes with particular suffixes which can only be viewed as either a punctual event or an iterative process.

What has made semelfactives particularly interesting for psycholinguistic investigation is the fact that when combined with a durative adverbial or in the progressive they are most naturally interpreted with an iterative reading. Consider (28).

(28) a. John coughed all night.
   b. John was coughing.

Here, John was coughing many times and coughing is understood as a durative activity – producing a number of coughs in a row. Tests showing the linguistic behavior of semelfactives can be found in Table 2.1.

Achievements (eventualities of type \(-\_\,-, e, f, \_\))

The last aspectual class, achievements, also cover instantaneous events. But in contrast with semelfactives, achievements change the state of the world. This is captured by the EC representation (29b) of sentence (29a). Before the reach event happened John wasn't at the top, but after the event he is.

(29) a. John reached the top.
   b. John reach-the top

| 1. Initially(~be-at-the-top) |
| 2. Initiates(reach(john,top), be-at-the-top, t) |

7. The first formula, Initially (~be-at-the-top), is strictly speaking unnecessary, because at the beginning of the scenario at time 0 it can be derived via negation as failure that HoldsAt(be-at-top, 0) is false. To do so, however, HoldsAt has to be defined as truth predicate which requires rather advanced logical machinery (e.g. EC uses the calculus by Feferman (1984)). To keep things simple, I refrain from doing so but instead explicitly introduce the initial situation of the scenario.
Achievements show a unique pattern with respect to compatibility with different kinds of adverbials. This is shown in (30).

(30) a. Five minutes ago, John reached the top.

b. *For five minutes, John reached the top.

In (30a) the achievement appears with a locating adverbial which states that an event happened five minutes before utterance time. This event can be identified with the reach event. (30b) is semantically anomalous because a durative adverbial cannot be said to hold for a punctual event. (31) shows, however, that an achievement modified by a durative adverbial becomes acceptable if it is in the progressive.

(31) For five minutes, John was reaching the top.

Why is (31) acceptable? The progressive coerces the achievement into an accomplishment which allows modification by a durative adverbial. (31) shows that the different Aktionsarten are intimately connected with grammatical aspect and can be shifted from one category into the other. These transitions between aspectual classes make the classification rather difficult and call for a semantic theory like EC that can deal with default interpretations and their reinterpretation due to aspectual coercion.

2.2.3 The classification of Aktionsart

To investigate the processing of Aktionsart it is crucial to determine the aspectual class to which a given expression belongs. I will follow Dowty (1979, p. 60) and his classification criteria. One has to be aware, however, that none of these tests works without exception because almost always there is the possibility to coerce the aspectual meaning. For example, John runs expresses an activity. However, it is also possible to interpret the sentence as expressing a habitual state of John. Dowty's tests will be used to determine the default interpretation of linguistic expressions when they are uttered out of the blue. The classification criteria are summarized in Table 2.1. As can be seen, the tests often do not allow to distinguish one category from all the others (as does e.g. the for-adverbial-test which distinguishes achievements from the other classes), but identify subsets of the categories (e.g. the in-adverbial-test distinguishes telic from atelic Aktionsarten). I have added a fifth aspectual category, the class of semelfactives.

In addition to Dowty's tests, I included the herum-test which exploits the contrast in acceptability between Hans joggte im Park herum (Hans was jogging in the park) versus *Hans joggte zwei Kilometer herum (Hans jogged two Kilometers around). This test is special for German and will be used to identify unambiguous activities since only these can be combined with herum. The only exception are semelfactives which are acceptable with herum, but are automatically coerced into an iterative reading. This is shown by Hans nieste im Klassenzimmer herum (Hans sneezed in the classroom around) which can only be interpreted as Hans sneezing more than once.

Table 2.1. Aspectual tests (adopted from Dowty (1979, p. 60))

<table>
<thead>
<tr>
<th>Criterion</th>
<th>State</th>
<th>Activity</th>
<th>Accomplish</th>
<th>Achievement</th>
<th>Semelfactive</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. meets non-stative tests</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
<td>if agentive: ok</td>
<td>if agentive: ok</td>
</tr>
<tr>
<td>2. has habitual interpretation in simple present</td>
<td>ok</td>
<td>ok</td>
<td>ok</td>
<td>bad</td>
<td>ok (iteratively)</td>
</tr>
<tr>
<td>3. φ for x time; spend n time φ-ing</td>
<td>bad</td>
<td>bad</td>
<td>ok</td>
<td>ok</td>
<td>bad</td>
</tr>
<tr>
<td>4. φ in x time; take x time to φ</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>d.n.a.</td>
<td>no</td>
</tr>
<tr>
<td>5. for x time entails at all times during that time</td>
<td>d.n.a.</td>
<td>yes</td>
<td>no</td>
<td>d.n.a.</td>
<td>d.n.a.</td>
</tr>
<tr>
<td>6. x is φ-ing entails x has φ-ed</td>
<td>ok</td>
<td>ok</td>
<td>ok</td>
<td>bad</td>
<td>ok</td>
</tr>
<tr>
<td>7. complement of stop</td>
<td>bad</td>
<td>bad</td>
<td>ok</td>
<td>bad</td>
<td>bad</td>
</tr>
<tr>
<td>8. complement of finish</td>
<td>d.n.a.</td>
<td>d.n.a.</td>
<td>yes</td>
<td>no</td>
<td>d.n.a.</td>
</tr>
<tr>
<td>9. ambiguity with almost</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>10. x φ-ed in x time entails x was φ-ing during that time</td>
<td>d.n.a.</td>
<td>d.n.a.</td>
<td>yes</td>
<td>no</td>
<td>d.n.a.</td>
</tr>
<tr>
<td>11. occurs with studiously, attentively, carefully, etc.</td>
<td>bad</td>
<td>ok</td>
<td>ok</td>
<td>bad</td>
<td>bad</td>
</tr>
<tr>
<td>12. can be combined with herum (&quot;around&quot;)</td>
<td>bad</td>
<td>ok</td>
<td>bad</td>
<td>bad</td>
<td>bad</td>
</tr>
</tbody>
</table>

note: ok = the sentence is grammatical, semantically normal; bad = the sentence is ungrammatical semantically ill-formed; d.n.a. = the test does not apply this category. φ stands for a VP (e.g. reach- the top).

The non-stative tests are a set of environments where only non-statives can appear. First, only non-statives occur in the progressive (I am *knowing/telling the answer). Second, only non-statives occur as arguments of force and persuade (I forced Joe to *know/tell the answer). Third, only non-statives can occur as imperatives (*know/tell the answer!). Finally, only non-statives occur in pseudo-cleft constructions (What Joe did was *know/tell the answer). Because states will play almost no role in my experiments, the stative tests are grouped together.
2.2.4 The composition of Aktionsart

Although Vendler classified different Aktionsarten of verbs it is well known that Aktionsart is a property of at least the verb plus its arguments (e.g. Verkuyl (1972)). Since language can in principle express infinitely many verb-argument-structures we need a compositional derivation of the Aktionsart of complex phrases. I will use three lexical features \([\pm \text{ADDTO}], [\pm \text{DUR}]\) and \([\pm \text{SQA}]\) that allow to derive Aktionsart in a compositionally transparent way. This is inspired by Verkuyl (Verkuyl 1972; Verkuyl 1993; Verkuyl 2005) who analyzes the composition of Aktionsart using the semantic features \([\pm \text{ADDTO}]\) and \([\pm \text{SQA}]\). Of course, there are other formal semantic accounts that would allow us to derive the Aktionsart in a compositional way drawing on semantically well defined mereological notions (see e.g. Bach (1986) or Krifka (1998)).

The feature \([\pm \text{ADDTO}]\) expresses if a verb introduces a path and can be directly read off from the lexicon. Take for instance walk which expresses changing spatial location over time. The path encoded by walk consists of pairs \((t, \text{place}(t))\) where \(t\) is some time and \(\text{place}(t)\) is the location at time \(t\). The term location is taken in a very broad sense, covering all sorts of semantic fields (position, color, possession, etc.). Both discrete and continuous changes in location receive a positive ADDTO feature. ADDTO receives a plus value just in case the denotation of the verb contains a location at a time \(t_i\) after \(t_0\) that differs in numerical value from \(\text{place}(t_0)\).

Activities, accomplishments and achievements all involve paths since they express a change in location over time. States have a negative ADDTO-feature since they are conceptualized to involve no change. Figure 2.1 illustrates the continuous path of walk and the discrete path introduced by reach while the stative verb know doesn't introduce any dynamic at all.

Achievements and semelfactives are conceptualized as punctual events while activities, accomplishments and states are durative in nature. That is, only the latter extend over time. By taking into account a durativity feature we can further distinguish between different kinds of verbs. For instance in Figure 2.1 the verbs walk and know receive a \([+\text{dur}]\) value while reach is \([-\text{dur}]\). DUR receives a plus value just in case all subintervals (of appropriate size) of the interval for which the verb \(P\) holds are also instances of \(P\). If walk holds for an interval \(I\) then it also holds for all the subintervals of \(I\) which are big enough to still count as walking. The same is true for the stative verb love. By contrast, not all the subintervals of reach can be characterized as instances of reaching thus reach has to be classified as \([-\text{dur}]\).

The specified quantity (SQA) feature expresses semantic properties of the arguments and has been introduced by Verkuyl (for a formalization see Verkuyl (1993)) who treats it within generalized quantifier theory. A \([+\text{SQA}]\)-NP pertains to something discernible that can be separated from other things and as soon as this can be done, one may count or measure. An argument receives a positive SQA feature if it is bounded, which means that it is able to impose an upper limit on a path (if it is an internal argument) or an event (if it is the external argument). Table 2.2 shows all possible combinations of features and the resulting Aktionsarten. It illustrates the plus principle: the only feature combination that yields an accomplishment is the one where all features carry a plus value.
Table 2.2. The standard Aktionsarten and their feature distribution illustrated for transitive verbs. \( \theta_1 \) marks the theta role of the internal argument and \( \theta_2 \) the theta role of the external argument.

<table>
<thead>
<tr>
<th>Aktionsart</th>
<th>ADDTO</th>
<th>DUR</th>
<th>SQA(( \theta_1 ))</th>
<th>SQA(( \theta_2 ))</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>State</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>John knows a poem</td>
</tr>
<tr>
<td>State</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>John knows poems</td>
</tr>
<tr>
<td>State</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>Children know a poem</td>
</tr>
<tr>
<td>State</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Children know poems</td>
</tr>
<tr>
<td>Activity</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>John ate sweets</td>
</tr>
<tr>
<td>Activity</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>Children ate a chocolate bar</td>
</tr>
<tr>
<td>Activity</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>Children ate sweets</td>
</tr>
<tr>
<td>Activity</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>Demonstrators kicked a policeman</td>
</tr>
<tr>
<td>Activity</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>A demonstrator kicked policemen</td>
</tr>
<tr>
<td>Activity</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Demonstrators kicked policemen</td>
</tr>
<tr>
<td>Activity</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>John won prizes</td>
</tr>
<tr>
<td>Activity</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>Athletes won a prize</td>
</tr>
<tr>
<td>Activity</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Athletes won prizes</td>
</tr>
<tr>
<td>Accompl.</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>John ate a chocolate bar</td>
</tr>
<tr>
<td>Achiev.</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>John won a prize</td>
</tr>
<tr>
<td>Semelfac.</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>A demonstrator kicked a policeman</td>
</tr>
</tbody>
</table>

Thus, prima facie (32a) and (32b) provide counterexamples to the feature analysis of Aktionsart. Let us have a closer look at them.

Considering sentences like (32a) Verkuyl (1993) argues that push-verbs need a [+SQA] goal-argument (like to the store) to be interpreted as accomplishments. In (32a) no such argument is overtly expressed in the sentence leading to a [−SQA]-value of an implicitly realized argument. The internal argument the cart is not of the right semantic type to impose an upper limit on the spatial path introduced by the verb. The example shows that the determination of Aktionsart relies on semantic properties like the location imposed by the verb.

Another problem is raised by (32b). Again, the feature configuration predicts an accomplishment, but (32b) expresses an activity. Furthermore, in (32b) it is not possible to explain the activity reading on the basis of an implicit argument. But a closer look at the lexical semantics of the verb might provide a clue. To entertain someone is a psychological verb of the amuse class (Levin 1993). These verbs do not imply that the experiencer is in the state of being entertained. It thus expresses a non-stative, durative action without an incremental theme. These verbs aren’t covered by the feature system introduced here, but since they do not play a role in the present work I will put these issues aside.

2.2.5 Hierarchical event structures

So far, I have introduced a typology of Aktionsart and have shown how Aktionsart can be determined on the basis of the lexical properties of the verb-argument-structure. EC allowed us to formalize the semantics of the different situation types. We are now in a position to come back to the issues that have been raised in connection with the psychological studies on events. First, we need to represent events in hierarchical structures that reflect different degrees of granularity. Second, we have to connect speaking and thinking about events to planning. The latter will be shown in the next section when we introduce the computational machinery of EC. The former is the topic of the present section.

Up to now, all the examples of events have been instantaneous (points like cough or reach and the start and finish events used in the scenarios). These punctual events are close to the bottom layer of an event hierarchy. A complex eventuality like an accomplishment can also be construed as a single event. This is shown in (33).

(33) Five years ago, John wrote his first novel. After that event his life changed completely.

The anaphor that event in the second sentence of (33) refers to the whole event expressed by the accomplishment John write his first novel. In some fashion, the whole eventuality has been packed into a single unit. The opposition of viewing the internal makeup of a situation or viewing it as a whole complete is reflected in linguistics by the notions of imperfective and perfective aspect. A classic description comes from Comrie (1976) who investigates aspect in a typological variety of languages:

"Aspects are different ways of viewing the internal temporal constituency of a situation [p. 3]; [...] Perfectivity indicates the view of the situation as a 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. [p. 16]"

EC rests upon this conception and provides a formalization. Imperfective aspect in the progressive sentence John was writing his first novel implies no completion of the goal and views the situation from an internal perspective. This can be easily formalized using the apparatus we already have. The accomplishment scenario contains the
activity fluent $\text{write}$ which holds as long as the event is ongoing. For the moment, let's represent the past progressive as

$$(34) \quad \text{HoldsAt}(\text{write}, t), \ t < \text{now}$$

which states that writing was going on at some time $t$ in the past.\footnote{This is not the proper way to deal with tense, however. The example is for illustrative purposes only. Tense will be introduced when the computational machinery is described in the next section.} It can easily be seen that by default this will lead to the completion of the novel. If, however, information about a terminating event is added to the representation this inference has to be withdrawn.

For perfective aspect a tool is needed to represent the hierarchical nature of events. EC utilizes hierarchical planning which provides a way to fuse the various elements of an eventuality into a single event. Consider the formal representation (35) of the perfective accomplishment John wrote his first novel (leaving tense aside).\footnote{Note that the rule is stated in logic programming with the conditional $\rightarrow$ having a non-classical interpretation. It is equivalent to a biconditional in the completion of the program. The semantics of the conditional is provided in Appendix A.}

$$(35) \quad \text{Happens}(\text{start}_{\text{write}}, s) \land \text{Happens}(\text{finish}_{\text{write}}, r) \land s < t \leq r \rightarrow \text{Happens}(e_1, t)$$

$e_1$ ("the writing of John's first novel") happened if there is a fluent $\text{write}$ which started at some time $s$ and was finished at some later time $r$. The definition of the temporal profile only uses ingredients from the accomplishment scenario. $\text{Happens}(e, t)$ will be false for all $t$ if the $\text{start}_{\text{write}}$ or $\text{finish}_{\text{write}}$ events do not happen, that is $e_1$ must contain all the events of the accomplishment. This explains why perfective accomplishments yield a contradiction when continued with but $\text{did not finish}$. Note that (35) is less informative than the original accomplishment scenario. Since it contains no $\text{Trajectory}$ predicate it abstracts away from the exact time course of the writing activity.

Let's move on and construct even higher levels of the event hierarchy. Assume we want to represent the complex event $e'$ expressed in the little discourse in (36).

$$(36) \quad \text{John wrote his first novel and sent it to the editor.}$$

$e'$ comprises two events, $e_1$ already defined in (35) and a second event $e_2$ (= John sent the novel to the editor) which can be defined analogously to (35). The complex event $e'$ is formalized in (37).

$$(37) \quad \text{Happens}(e_2, t) \land \text{Happens}(e_1, t) \land s < t \rightarrow \text{Happens}(e', t)$$

Applying hierarchical planning recursively allows us to formulate a statement that represents a whole discourse in terms of a plan-goal structure at a very coarse grain-size. This rather abstract event-structure can be decomposed all the way down to the concrete events that are at the bottom layer of the hierarchy. Hierarchical planning elegantly models the levels of granularity discussed in the psychological studies on event cognition in Section 2.1.1.

2.2.6 A glimpse at the computational machinery

In this section I will sketch how EC computes the temporal profile of a situation. The reader is encouraged to consult Hamm & van Lambalgen (2005) for the complete formal version (see also Appendix A) for the semantics of propositional logic programming. Given a sentence or even whole discourse, EC computes a situation model that makes the discourse true. In contrast to classical logic EC doesn't have to deal with a potentially infinite number of models all satisfying the discourse, instead it always computes a unique model which is the minimal model of the discourse.\footnote{This only holds if the discourse isn't contradictory. If it is, then no model gets computed.} Minimality refers to the requirement that the model contains only those occurrences of events forced to be there by the discourse and the axioms. Logic programming fulfills minimality by incorporating $\text{negation as failure}$, that is everything one has no information about is assumed to be false. As a result, in a discourse model no unforeseen events are allowed to occur.

In EC, a sentence serves as an instruction to update the discourse model in such a way that the sentence is true. To do this, we have to distinguish two kinds of information. Universally quantified information appears in the scenario and in the axioms. The scenario contains lexical information about the temporal profile of the situation described by the sentence. By contrast, tense introduces existentially quantified information, for instance that there is a time $t$ in the past and fluent $f$ is true at $t$: $\text{HoldAt}(f, t), t < \text{now}$. Existential information cannot be handled by merely adding the latter statement to the scenario.\footnote{Doing so would yield the incorrect result that for all time instants (in the past) $f$ is true.} Instead the statement $\text{HoldAt}(f, t), t < \text{now}$ triggers abductive reasoning using the axioms of EC and the scenario information. In logic programming, this sort of reasoning is carried out by a derivation procedure called resolution (more specifically SLDNF resolution, see e.g. Nienhuys-Cheng & de Wolf (1997)). It starts with the formula we want to make true in the discourse model. In EC, this top query is introduced via an $\text{integrity constraint}$. 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 the updated model. To get an impression of how this works let’s go through the resolution of sentence (38a) in present perfect with the scenario (38b) and the integrity constraint (38c).

(38) a. John has reached the top.
    b. John reached the top
    1. Initially(¬ on top)
    2. Initiates(reach, on top, t)
    c. ¬HoldsAt(on top, now) succeeds

Resolution starts with the integrity constraint. HoldsAt(on top, now) can be resolved using Axiom 3 and we get the new query (39).

(39) ¬HoldsAt(on top, now) Axiom 3
    ¬Happens(reach, t′), Initiates(reach, on top, t′),
    ¬Clipped(t′, on top, t) t′ < t, t < now

Initiates(reach, on top, t′) is a fact stated in the scenario, but Happens(reach, t′) has to be added to the discourse model under construction. Let’s turn to the query ¬Clipped(t′, on top, t). A negated formula is resolved by showing that its positive counterpart fails. This is done by adding a subtree for the positive formula to the resolution tree and showing that it fails. In our example no t-relevant events are assumed to happen after t′ making the subtree query fail; hence ¬Clipped(t′, on top, t) is true. This step is shown in (40).

(40) ¬Happens(reach, t′), Initiates(reach, on top, t′),
    ¬Clipped(t′, on top, now), t′ < now
    ¬Happens(e, t′), t′ < now, (Terminates(e, on top, t′),
    ¬Releases(e, on top, t′))
    ¬Clipped(t′, on top, now) Axiom 5
    ¬Happens(e, t′), t′ < now, (Terminates(e, on top, t′),
    ¬Releases(e, on top, t′))
    ¬Happens(e, t′), t′ < now, (Terminates(e, on top, t′),
    ¬Releases(e, on top, t′))
    ¬negation as failure

The last query cannot be further resolved, so we are done. The discourse model is updated with a reach event at time t′ that happened before now. The on top fluent states that John was not on the top during the time interval up to and including event time t′ which abuts a later time interval for which the fluent on top is true. This time interval holds up to now capturing the present relevance of the English present perfect.

The resolution of (38a) can be used to show the non-monotonic nature of the update process. Without further information we get the default interpretation that before time t John has never been on the top. If, however, the sentence is continued with once again the default interpretation is abandoned and on top will have to be constructed with a more complex temporal profile.

EC is a framework that nicely fits the requirements imposed by the studies from cognitive psychology on a linguistic theory of events. We have seen how the temporal profile of a minimal situation or Aktionsart is computed and updated with additional information. We have also seen how EC encodes event hierarchies up to the discourse level. The next section will be about work on aspectual reinterpretation, again both from a formal semantic and from a psychological perspective.

2.3 Coercion

Up to this point we have been dealing with the default case in aspectual interpretation: all the constituent parts of the Aktionsart fit together and interpretation works smoothly in a strictly compositional fashion. Nonetheless, there may be a clash between expressions of different semantic sorts. Still, the interpretational system often succeeds in arriving at a coherent interpretation. This phenomenon is called coercion. Coercion isn’t restricted to the aspectual domain, but also shows up in reinterpreting a noun in so-called complement coercion. The latter has attracted a fair amount of attention in psycholinguistics (e.g. McElree, et al. (2001)). The contrast in (41) provides an example.

(41) a. The author began writing the book.
    b. The student began reading the book.
    c. The author began the book.
    d. The student began the book.

In opposition to (41a) and (41b), (41c) and (41d) lack relevant pieces of information. Begin requires an argument which denotes an event. (42) shows that a physical object is inappropriate.

(42) a. John began to row the boat.
    b. *John began the boat.

When a comprehender encounters the book in (41c) and (41d) he has to add information in order to impose an event interpretation on it. In particular, he must determine what kind of activity the agent was engaged in. For instance, was it writing the book, reading it, or learning it by heart? Note that the preferred interpretation
differs between (41c) and (41d). This difference shows us that the chosen interpretation depends on stereotypical knowledge; while authors typically write books, students most probably read or study them.

Psycholinguistic investigations have shown that coercion phenomena cause processing difficulty during reading (e.g. McElree, et al. (2001)). But before coming to these studies I will present semantic theories on reinterpretation to gain a better understanding of what processes may underly the computation of coerced meanings.

### 2.3.1 Semantic theories on coercion

In this section, I will compare classes of semantic coercion analyses and will relate them to approaches of type coercion in general. The work of Moens (1987) and Moens & Steedman (1988) who introduced the term coercion into the literature will serve as a starting point. The theories that came after Moens & Steedman (1988) can roughly be classified into three categories:

- **Operator-Based Accounts** assume that coercion is triggered by a type-mismatch between two expressions A and B that have to be composed. Let expression A be a functor of type \( \beta, \gamma \) (meaning that A takes an argument of type \( \beta \) and outputs a value of type \( \gamma \)); let expression B be an argument of type \( \alpha \) with \( \alpha \neq \beta \). A type-mismatch is a constellation \([A_{\beta,\gamma}, B_{\alpha}]\). It can be resolved by inserting a type-shifting operator \( \text{OP}_{\alpha,\beta} \) from type \( \alpha \) to type \( \beta \) into the representation yielding \([A_{\beta,\gamma}, \text{OP}_{\alpha,\beta}(B_{\alpha})]\). The latter can be interpreted compositionally.

The type-shifting operator can appear in the semantic representation without having a syntactic counterpart. Alternatively, it can be accompanied by a corresponding reanalysis of the syntactic structure. The latter alternative is particularly suited for a lexical decomposition account of complex predicates in syntax.

- **Underspecification Accounts** have been developed to represent ambiguities in an economical and elegant way. They use a two-step procedure. First, an underspecified representation is computed. This representation – the semantic representation proper – leaves everything open which needs further disambiguation. In a second step, the underspecified representation is transformed into a fully specified representation using pragmatic information from discourse context, conceptual knowledge and so forth.

A coerced sentence and its non-coerced counterpart share the same underspecified representation. Only the second step is different. While the non-coerced sentence is disambiguated by eliminating all potential coercion sites in the semantic representation, the coerced sentence is disambiguated by inserting coercion operators to fully specify the meaning.

- **Planning Accounts** make use of planning formalisms which have been developed in artificial intelligence. According to these theories, the computation of the temporal profile for a piece of discourse can be viewed as coming up with a plan that permits a non-monotonic derivation of a goal state from an initial state. We have already made acquaintance with one planning account, the Event Calculus by Hamm & Van Lambalgen (2005). Coercion is characterized by shifts in meaning that result from the update of a plan when integrating new information. Planning accounts allow us to differentiate substantially different types of coercion by the computations that are involved. Because situation models are computed in a computationally tractable way, planning accounts enable us to derive fine-grained predictions concerning the processing difficulty of coerced meanings.

---

**Complement coercion and the generative lexicon**

Psycholinguistic studies have mainly focussed on complement coercion and discussed their findings in the framework of the generative lexicon by Pustejovsky (1995). I will therefore first provide a brief introduction to Pustejovsky’s account. He develops a theory of lexical meaning with highly structured lexical entries. These entries are intended to reflect the deeper conceptual structures in the cognitive system. A book, for example, isn’t a physical object the same way a stone is, but contains information, is made for reading, was created by being written etc. The complete conceptual structure is stated in what Pustejovsky termed *qualia structure*. Let’s have a look at the lexical entry of book in (43).

\[
\begin{align*}
\text{BOOK} & = [\text{ARG1} = x:info] \\
& \quad [\text{ARG2} = y:physobj] \\
& \quad [\text{X.Y-LCP} = \text{hold}(x,y)] \\
& \quad [\text{TELIC} = \text{read}(e1, reader, x,y)] \\
& \quad [\text{AGENT} = \text{write}(e2, author, x,y)]
\end{align*}
\]

*Book* has two primitive word senses: a physical object denotation (e.g. *the book is on the table*) and an information denotation (e.g. *the book is interesting*). These two senses and their combination, the dotted type \( x,y \), form a cluster which Pustejovsky refers to as a *lexical conceptual paradigm* (lcp). It captures the polysemy of the noun by relating it to the whole type hierarchy in (44).

\[
\begin{align*}
\text{x: information} & \quad \text{y: physical object} \\
\text{x.y: print matter} & \quad \text{book}
\end{align*}
\]

The qualia of an object can be seen as the initial points from which to construct interpretations that would otherwise be ill-formed. The *formal* quale defines the relation
between the arguments of the different types in a complex type. A book is a sort of container – its physical manifestation holds symbolic information. The Telic quale expresses the purpose of a book: books (in the sense of print matter) are for reading. Finally, the Agentive quale expresses how books come into being: they are written by authors. In complement coercion, the qualia structure is used to construct new lexical meanings in a principled way. Let’s consider the standard complement coercion example repeated in (45).

(45) The author began the book.

Begin selects an event, but the book doesn't have an event denotation. According to Pustejovsky, the event reading has to be reconstructed from the qualia structure. In principle there are two qualia that have an event argument allowing for possible reinterpretations. Because one of them, the Agentive quale, corresponds to the typical activity authors engage in, the book will be coerced into a writing event rather than a reading event.

A problem with Pustejovsky's theory is that it both over- and undergenerates coerced meanings. In (46a) the intended meaning cannot be construed on the basis of the qualia structure.12 The lexical entry of novel probably doesn't contain information that novels are lectured on in literary studies. Thus, the theory cannot generate the meaning of the sentence. On the other hand, Pustejovsky's account overgenerates coerced meanings: the telic quale of cigarette definitely refers to a smoking event, but without explicit mention of the infinitive to smoke, (46b) sounds rather odd.

(46) a. The professor of literary studies began (to lecture on) the novel.
   b. The heavy smoker began (to smoke) the cigarette.

Ignoring these problems, the generative lexicon offers a flexible lexical semantics that takes into account the context in which a lexical item appears. Furthermore, it can be used to derive predictions about the processing of complement coercion. First, coercion is expected to cause more difficulty than a non-coerced sentence because the meaning has to be reconstructed from the qualia structure. Second, the theory predicts that difficulty should be solely due to lexical processing. Third, the set of possible reinterpretation candidates should be rather restricted. We will come to some of these points when reviewing the psycholinguistic literature on complement coercion.

**Aspectual coercion – The transition network**

In their seminal paper on aspectual reinterpretation, Moens & Steedman (1988) show that the aspectual classes are not fixed, but can be easily transformed into one another by means of linguistic or extralinguistic context. In principle, a verb-phrase can basically be shifted into each of the Vendlerian classes. For example, in (47) an achievement (47a) is shifted into a state (47b) and an accomplishment (47c).

(47) a. John won the race.
   b. For many years, John won the race.
   c. John was winning the race in fifty minutes.

Moens & Steedman (1988) introduced the term coercion into the semantic literature which they adopted from programming languages. In their framework, the five aspectual categories appear as the states of a finite-state automaton and coercion is modeled in terms of transitions between them. Diagram 2.2 shows their classification plus the network of possible transitions.

---

To play the "Minute Waltz" introduces an activity (= 48-i). When combined with in less than 60 s it is shifted into an accomplishment (= 48-ii). Because for more than an hour requires an activity, the accomplishment has to be type-shifted again. Only one route through the transition network allows this: first the accomplishment is reinterpreted as a point. Then it is iterated into an activity. The resulting accomplishment in (48-iii) describes the iterated activity of playing the "Minute Waltz" in less than sixty seconds as lasting for more than an hour. In (48-iv) the adverbial it took me two days behaves like an in-adverbial and needs an accomplishment. Although the derived accomplishment is of the right type, a conflict emerges between the length of time of the two adverbials for more than an hour and it took me two days. Thus, another route through the network is needed; the accomplishment has to be interpreted as an achievement first – to succeed in playing the MW in less than 60 s for more than 1 h. Second, turning this achievement back into an accomplishment a preparatory process is added. This process must be inferred. A plausible candidate is practicing, that is playing the MW repeatedly until the quality of the performance has reached the criterion specified by the culminating. The final interpretation of sentence (48) is that the speaker had to practice for two days until he acquired the skill to repeatedly play the MW for more than an hour in such a way that each of these performances was shorter than one minute.

The Moens and Steedman system of aspectual coercion has two very interesting features. First, it offers a complete and comprehensive system of the kind of coercions that are possible. Second, the different transitions are discussed in terms of different processes, for instance, iteration versus adding a suitable preparation. An interesting question for psycholinguistic investigation is whether all of these transformations cause processing difficulty and whether some of them are more demanding than others.

2.3.2 Operator-based accounts

Coercion poses a challenge to semantic theories because seemingly mismatching constituents can obviously be integrated into a coherent interpretation. One way to achieve a formal semantic analysis of coercion is to assume that certain type-shifting operators are inserted into the representation whenever a type-mismatch occurs. I termed this kind of theoretical models operator-based accounts. According to them the representational system can work in one of two alternative modes. One is surface interpretation where everything works fine and smoothly. The other one is coercion where local semantic mismatch has to be resolved by plugging an operator into the semantic representation. In principle, operator-based accounts come in two flavors: either coercion only affects the semantic representation or it is reflected by changes of the syntactic structure, too.

**Introducing a semantic operator**

I will exemplify the first kind of operator based account by describing the coercion theory of de Swart (1998). In her theory, a coercion operator is inserted into the semantic representation if and only if a type-mismatch occurs. Compared to a non coercing sentence coerced sentences are semantically more complex, but syntactically they are the same.

De Swart presents an approach to aspectual coercion within Discourse Representation Theory (DRT, see Kamp & Reyle (1993), a brief introduction to DRT is also provided in Appendix B), one of the most influential current approaches to the semantics of natural language. She derives coerced meanings by plugging operators into the Discourse Representation Structure (DRS) which shift the type of the mismatching eventuality. These operators have no syntactic equivalent. Let's analyze the two English progressive Examples (49) which are structurally identical, but (49b) requires coercion, whereas (49a) does not.

\begin{enumerate}
\item a. John is pushing a cart.
\item b. John is loving that woman.
\end{enumerate}

The progressive takes an activity argument and outputs a state. The aspect operator PROG is the semantic counterpart of be V-ing in morpho-syntax. (50) is the semantic representation of (49a).

\begin{center}
\begin{tabular}{|c|}
\hline
\( s \) : PROG \\
\hline
\end{tabular}
\end{center}

The outer DRS expresses that the state of John is located within the reference time \( t \). The sentence is in the present, therefore the utterance time \( n (=\text{now}) \) is located within \( t \). The two discourse referents \( x \) and \( y \) represent John and the cart respectively. The progressive operator PROG is embedded inside a sub-DRS which itself embeds another DRS which introduces a process \( p \). The latter corresponds to the sentence radical John push a cart. All the parts of the DRS represent overtly realized morphosyntactic information, namely present tense, progressive aspect and the verb with the subject plus the
direct object. There is no type mismatch in deriving the meaning. Let’s compare it with the DRS (51) representing the coerced sentence (49b):

(51) \[
\begin{array}{c}
\text{nt}\text{s}\text{sx}\text{y} \\
\end{array}
\]

John (s) that woman (y)

\[
\begin{array}{c}
t \subseteq t \\
s = t \\
\text{d: PROG} \\
\text{s: Csd} \\
\end{array}
\]

\[
\begin{array}{c}
\text{d'} \\
\text{s': love (x, y)} \\
\end{array}
\]

\[
\text{Csd}
\]

Love is a stative verb. Recall that the progressive was the most important test to distinguish statives from non-statives. Why can a stative verb be progressivized in this example at all? The answer is that in (49b) the state John love-that woman is coerced into a non-stative eventuality. This is exactly what is captured by the representation in (51). Two steps are necessary:

First, the inner DRS introduces a state s’ which has to be turned into a dynamic eventuality d. This is done by making the semantic representation more complex and inserting a yet abstract type-shift operator Csd which turns a state into a non-stative.

The second step is interpreting Csd as a function that maps a concrete state description onto an activity. Only if the abstract operator Csd can be made concrete, is coercion successful. The right kind of non-stative eventuality has to be inferred on pragmatic grounds, for instance, loving somebody to a certain degree. This interpretation becomes even more natural if the context contains the right kind of information. This is shown in (52).

(52) John is loving that woman more and more, the more he gets to know her.

Based on this analysis we can derive interesting predictions about the processing of coerced sentences. First, the two-step procedure I have just described should be reflected by two consecutive processing operations. One is a semantic type-shift triggered by the occurrence of aspectual mismatch; the other is the actual repair in which the right kind of operator has to be chosen. Second, the theory predicts that context information can help to select the appropriate type-shift operator. It therefore can eliminate difficulty having to do with the second step. Context shouldn’t, however, have an influence on the preceding purely semantic operation.

Operators that are also realized in syntax

A syntacto-semantic account of coercion assumes that coercion affects both the semantic and the syntactic representation. On the basis of morphosyntactic evidence from some Western Malayo-Polynesian languages like Tagalog and Malagasy it has been claimed that complex predicates are decomposed in syntax (Travis 2003). The basic idea of lexical decomposition is to express complex predicates by more primitive or basic ones. The classical example is kill which has been analyzed as cause somebody to become not alive. Primitive predicates like become have a long history. Dowty (1979), for instance, uses cause, become, do and be as basic primitives to express the four Vendlerian classes. A syntacto-semantic decomposition account assumes predicate decomposition not only in the semantics, but also in the syntax. For instance, under the assumption of syntacto-semantic decomposition (53a) has a structural representation along the lines of (53b) (see e.g. McCawley (1968)).

(53) a. John kills Mary.
   b. [John cause[become[not[alive[Mary]]]]]

What happens if we take coercion into account? Let’s look at an example where coercion leads to the enrichment of the underlying representation. Assuming syntactic decomposition, an achievement like (54a) has the relatively simple structure (54b).

(54) a. John reached the top.
   b. [John[become[be[on the top]]]]
   c. John was reaching the top.
   d. [John[do x[cause[become[be[on the top]]]]]]

In (54c) the sentence is in the progressive and has the much richer decomposition in (54d). Rothstein (2004, p.136–139) defines the shift operation in (55) for achievements in the progressive. The shift inserts all the missing pieces of the decomposition analysis into the semantics.

(55) \[
\text{SHIFT}_{\text{achievement-achievement}}(\text{VP}):: \rangle
\]

The shift operator takes an achievement and makes it part (= e_1) of a new achievement. In addition, a preparatory process e_2 is added. The two atomic events e_1 and e_2 are summed (= e_1 \lor e_2) and form the accomplishment e which culminates in the
achievement. The preparatory process is left unspecified and has to be determined contextually. To yield a syntacto-semantic account, we just have to take this analysis one step further (which Rothstein doesn’t do!) not only assuming that the achievement is reconstructed semantically, but also that the shift operator introduces an additional layer in the phrase marker. Alternatively, the VP may be reanalyzed introducing additional structure corresponding to the unspecified activity.

If an account along these lines is correct, we can derive the following predictions about the processing consequences of aspectual coercion. To repair local aspectual mismatch, an appropriate type-shift operator has to be inserted into the semantic representation. Application of the type-shift operator has two consequences: The semantic representation will be updated by adding events to the semantic representation (or subtracting them, depending on the kind of coercion). Second, the syntactic representation will be reanalyzed by inserting (or deleting) syntactic structure that corresponds to the decomposition analysis of the coerced sentence. Thus, processing a sentence like (54c) should also have visible processing consequences that are due to syntactic reanalysis.

2.3.3 Underspecification

Underspecification formalisms were designed for the representation of semantically underspecified expressions, i.e. expressions whose meaning can only partially be determined by the semantic construction. In contrast to operator-based theories they assume that coerced and non-coerced sentences are semantically identical. A fully specified meaning is computed in a two-step procedure: In the first step, an underspecified representation is created. This representation – the semantic representation proper – leaves everything open which needs further disambiguation. In the second step, the underspecified representation is disambiguated on the basis of pragmatic information like discourse context and conceptual knowledge.

There are three main proponents of aspectual underspecification: Pulman (1997), Dölling (e.g. Dölling 2003a) and Egg (2005). Since their theories make different predictions I will present them separately.

Before starting to review the different underspecification accounts, a general remark may be in order. None of these theories was invented as a processing account and in case of a type mismatch coercion operators are modeled as an instantiation of a variable that was previously introduced into the semantic representation. When no type-mismatch occurs the coercion operator is instantiated as the identity mapping, whereas in case of a type mismatch coercion operators may recursively be applied. Which operator has to be chosen depends on world knowledge about the typical duration of events. For instance, John sneezed for three seconds is aspectually ambiguous; both an iterated reading and a slow motion reading are possible in this example.

From Pulman’s theory it follows that multiple coercions can occur at the same syntactic position. However, the experiments in the following chapters will provide evidence that this is not the case. If an achievement is combined with a for-adverbial (e.g. the reached the top for two hours) the sentence is judged to be semantically ill-formed. In principle, two coercions are needed to make sense of it. First, the achievement has to be coerced into an accomplishment and second, the accomplishment has to be shifted into an activity. Note that both coercions are unproblematic if they appear in isolation, but when they are combined the construction is not interpretable anymore. This is not a complexity issue, because if we add another “coercion site” to the sentence by putting it in the progressive it becomes acceptable: he was reaching the top for two hours. Thus, multiple coercion does not seem to be possible.

Dölling’s approach

Dölling’s work (Dölling 1995, 1997, 2001, 2003a) offers an in-depth analysis of coercion in the nominal and in the aspectual domain. The analysis is similar in both domains and aspectual and complement coercion are assumed to involve the same two steps. First the underspecified representation is computed in a strictly compositional fashion. In a second step the representation is specified bearing on world knowledge. In the following, I will only concentrate on aspectual coercion and present a sample analysis along the lines of Dölling (2003a). Consider the coerced VP in (56a) together with the corresponding underspecified semantic derivation tree in (56b).

(56) a. sneeze for ten minutes

b. sneeze for 10

- adverbial modifies

- shift

- no shift
Reinterpretation is anticipated by inserting an abstract reinterpretation template called shift. It is inserted between any functor-argument pair and is applied to arguments regardless of the properties of the functor they are combined with.\(^{14}\) In (56b) the shift operator has been applied twice. It has the following semantics:

\[
\langle \text{shift} \rangle = \lambda P.\lambda x. Q y[R(x, x) \wedge P(y)]
\]

where Q and C are either \(\exists\) and \(\wedge\) or \(\forall\) and \(\rightarrow\). R can be any two-place predicate; if R is instantiated as \(= (x, y)\) under existential quantification it will be abstracted away.

(58) is the underspecified representation of the modified VP in (56a). In principle, it has to be shifted once more.\(^{15}\) The shifted VP can then be modified by other adverbial modifiers. This is needed in a sentence with multiple coercions like John managed to sneeze for ten minutes in only one day.

\[
(\text{sneeze for 10 minutes}) =
\]

\[
\lambda x. Q y[R(x, x) \wedge C y] y x \wedge \exists z[(z, x) \rightarrow \text{for}(y) \geq 10 \text{ min}]
\]

Let’s have a look at the specification stage. At this step, possible specifications are generated by substituting the \(Q, C, R\) with “real” quantifiers, connectives and two-place predicates. (59) is a conceptually possible specification of (58) which encodes an iterative reading.CONST is a relation between processes and events and expresses that a process p is constituted by event particulars e.

\[
(\text{sneeze for 10 minutes repeatedly}) =
\]

\[
\lambda x. \forall y[\exists z[x, y] \rightarrow \text{sneeze}(y)] \wedge \exists z[(z, x) \rightarrow \text{for}(z) \geq 10 \text{ min}]
\]

\[
\lambda p. \forall e[\exists z[x, y] \rightarrow \text{sneeze}(e)] \wedge \forall e[\exists z[(z, x) \rightarrow \text{for}(p) \geq 10 \text{ min}]
\]

What are the implications for the processing of coerced sentences? In contrast with Pulman’s underspecification account Dölling introduces only one shift operator for each argument. As a result, only one coercion should be possible at once. However, since in Dölling’s theory everything depends on pragmatic specification we need to know more about possible specifications. What are the possible instantiations of the two-place predicates \(R\) in the coercion template? Are there any constraints imposing restrictions on the available readings?

Unfortunately, it is impossible to derive detailed predictions about the available meanings before the pragmatic disambiguation mechanism is spelled out in more detail. The same holds with respect to expectations about processing difficulty: Is coercion more difficult than processing non-coerced sentences? Are different specifications (e.g. iterative vs. habitual readings) equally difficult to process? To answer these questions the mechanisms underlying specification would have to be worked out in more detail.

**Egg’s approach**

There is an interesting fact about the available scope readings in (60).

\[
(60)\begin{align*}
\text{a.} & \quad \text{John began to read every novel.} \\
\text{b.} & \quad \text{John began every novel.}
\end{align*}
\]

(60a) has two scope readings. It can either mean that John began reading the complete set of salient novels or it can mean that for each novel, John has begun reading it. The former reading corresponds to wide scope of begin; the latter is the wide scope reading of every. Interestingly, (60b) is unequivocal in scope: only the wide scope reading of every is available in this example. This is totally unexpected in the generative lexicon of Pustejovsky (1995). Egg (2005) offers an underspecification account that unifies the representation schemes of aspectual and complement coercion with underspecified representations of scope ambiguities. Bringing these two phenomena together he is in a position to explain the scopal differences between (60a) and (60b). I will present Egg (2003)’s analysis of the two examples to provide an idea about how his account works.

First, he draws a semantic distinction between VPs that include a control verb and VPs with a transitive verb. Only control verbs can scope high while transitive verbs always scope low. This is shown in (61a) with begin as a control verb versus non-coerced, transitive begin in (61b).

\[
(61)\begin{align*}
\text{a.} & \quad \text{John began to start every fight.} \\
\text{b.} & \quad \text{John began every fight.}
\end{align*}
\]

The analysis is carried out in an underspecification formalism called *constraint language for lambda structures* (CLLS). It has three main ingredients: (1) fragments of \(\lambda\) terms, (2) so called holes \(\square\) which have to be further specified and (3) dominance relations between \(\lambda\)-fragments and holes. The latter are indicated by lines and can be read from the bottom to the top as place this fragment (somewhere) inside this hole. Scope ambiguities have a representation in which the scope bearing elements are located at the same hierarchical level. This yields their characteristic diamond shape. A scope ambiguity is exemplified in (62) which is the underspecified representation of the non-coerced control sentence (60a).
Choosing an operator is determined on pragmatic grounds and all the conceptually appropriate operators constitute the set of reinterpretation candidates. Having chosen an operator our example has exactly one solution for specification, namely the one provided in (66). Since this solution is an alphabetical variant of (63a), the two readings are equivalent.

\[
\lambda e. \forall y. (\text{novel}(y) \to \text{begin}(\text{john}, \lambda e. \text{read}(x, y)(e)))(e)
\]

Aspectual coercion receives a structurally similar analysis like complement coercion. Again, we have a type mismatch – this time a mismatch in aspectual class – which is repaired by bringing a coercion operator into the underspecified representation. Operators include iterate, add preparation and so forth.

Like the previous accounts, Egg (2003, 2005) doesn't provide a theory of how and when pragmatic information is brought into the specification process. It is thus difficult to derive predictions which can be tested experimentally. However, since complement coercion and aspectual coercion are analyzed the same way they should be processed in a similar fashion; if the proposed analysis is on the right track aspectual coercion and complement coercion should, for instance, elicit the same functional correlates in neurophysiological measures such as event related potentials.

To sum up, underspecification accounts assume a two step procedure: First an underspecified semantic representation is computed which buffers type conflicts. Type mismatch is only resolved in a second step based on conceptual information. The existing accounts are too vague concerning the specification stage to make clear predictions for processing. If, however, underspecification is understood in a way that the aspectual representation remains underspecified as long as no disambiguating information is provided we get much stronger predictions. A psycholinguistically more motivated notion of underspecification along these lines will be discussed in Section 2.3.8.

2.3.4 Planning accounts

Planning accounts view a sentence as a goal that has to be made true. The comprehender must come up with a plan that connects an initial state (the discourse model without the information of the sentence) with the goal state via a sequence of events. Once a plan can be made true we can proceed to the next sentence. This is provided we get much stronger predictions. A psycholinguistically more motivated notion of underspecification along these lines will be discussed in Section 2.3.8.

Planning and affordances – Steedman’s approach

How can a coercion sentence like (67a) corresponding to the goal in (67b) be made true? I will illustrate the computation of a coerced meaning with an example similar to one from Steedman (2005, p. 14).

Chapter 2. Events in cognitive psychology and linguistics

2nd proofs
The Processing of Events

48

The Processing of Events

very important to be very clear about the information in the database. ever, that many different computations may be possible to derive a given goal. It is thus derive predictions about processing difficulty in coercion. We have to be aware, how.
sures like these the theory naturally relates to language processing and can be used to
that have to be formulated in the course of the derivation. Based on complexity mea
computational complexity. For instance, one can simply count the number of subgoals
An interesting property of Steedman’s account is that the formalism offers a measure of
This yields the new subgoal (69) which is provably true applying fact (4) and unifying
The yet abstract activity Q can now be unified with climb(x, K2) relying on rule (3). This yields the new subgoal (71) which is provably true applying fact (4) and unifying x with John.
(71) in_progress(climb(x, K2))
An interesting property of Steedman’s account is that the formalism offers a measure of computational complexity. For instance, one can simply count the number of subgoals that have to be formulated in the course of the derivation. Based on complexity measures like these the theory naturally relates to language processing and can be used to derive predictions about processing difficulty in coercion. We have to be aware, however, that many different computations may be possible to derive a given goal. It is thus very important to be very clear about the information in the database.

2nd proofs

Another interesting feature of Steedman’s approach concerns the way in which relevant pieces of conceptual knowledge are selected, e.g. rule (3). He argues that affordances in the sense of Gibson (1979) play a crucial role in determining the relevant actions. Affordances are the interactions with the world that a perceived object mediates. A door “affords” egress and ingress, a mountain affords climbing and a hammer affords grasping and pounding. Steedman (2005) argues convincingly that without an affordance-based approach linguistic theory helplessly overgenerates meanings.

“Many lexically-governed coercions are derived from nouns. Thomason (1997, p. 820) points out that identifying the meaning of phrases like “hammer the metal flat” with that of “causing the metal to become flat by hitting it with a hammer” overgeneralize to models such as those in which the hammering merely signals to a third party that they should put the metal through a hydraulic press. He suggests an analysis paraphrasable as ”using a hammer in the normal way for metal to make the metal flat.” He points out that “the information about normalcy that is needed in such examples is exactly the sort of information that is needed for practical planning.” We have reason to link this observation to what Gibson (1979) called the “affordances” of objects – that is, the events made possible by objects such as hammers, such as beating metal, and the consequent states of those events, such as the metal in question being flat.” Steedman (2005, p. 15)

To sum up, Steedman’s account provides two necessary prerequisites to a processing theory of aspectual coercion. First, the notion of computational complexity can be defined by algorithmic properties of the derivation. This is especially important when comparing different kinds of coercion from the transition network with each other. Second, using an affordance-based pragmatics the interplay of semantics and conceptual knowledge can be properly modeled. The possible actions for an object are restricted to its set of affordances.

Coercion in EC

Like Steedman’s theory, EC addresses the computational complexity of coerced meanings. In the next chapter I will use EC as a processing theory and present sample derivations for different kinds of coercions spelled out in a combined DRT/EC framework. Here I will concentrate on the way EC differs from Steedman’s theory.

The major differences between Steedman (1997, 2002, 2005) and EC are the following. While Steedman’s theory explicitly provides axiomatized coercion operators (e.g. see rule (1) in 68), EC does coercion completely without operators. Instead, a grammatical construction such as the progressive directly “moves” an expression from one aspectual category to another.

In addition, Hamm & Van Lambalgen (2005) propose a classification of different kinds of coercion which I will make use of throughout the rest of the book. Recall the
definition of eventualities (definition 5) formalizing the event nucleus. In principle, there are three possible forms of operations to change an eventuality. The situation can be elaborated by adding some eventualities to it via additive coercion (e.g. activity \(\rightarrow\) accomplishment); parts of a scenario can be deleted via subtractive coercion (e.g. accomplishment \(\rightarrow\) activity); or a complete event nucleus can be shifted into a different Aktionsart in cross coercion. An example of the latter is when the temporal structure of a process is imposed on an event or a state (e.g. semelfactive \(\rightarrow\) activity). Sample derivations of all three kinds of coercion will be provided in the next chapter.

2.3.5 A short summary

In this section I have presented a coarse overview over the vast semantic literature on reinterpretation phenomena. I classified the existing theories into three categories. Operator-based accounts assume that in coercion the semantic representation is enriched by inserting a coercion operator. In underspecification theories type-conflict is buffered by underspecified representations. When specifying the underspecified representation coercion operators are inserted whenever they are needed. Specification heavily relies on pragmatic information which isn’t explicitly modeled in underspecification accounts. Finally, planning accounts model aspectual interpretation as coming up with a plan that connects the initial discourse context with the information stated in the current sentence. Within these theories coerced sentences can be characterized as being computationally more complex than sentences without coercion. In addition, Steedman’s use of affordances offers an elegant account of identifying and restricting the pragmatically relevant actions that can go into coerced meanings.

2.3.6 Psycholinguistic studies on coercion

Psycholinguistic research on coercion has focussed on complement coercion. Aspectual coercion has received less attention. Both have been used to challenge the principle of (strict) compositionality: 16

“The meaning of an expression is a function of the meaning of its parts and of the way they are syntactically combined.” (Partee 1984)

In what follows I will consider psycholinguistic and neurolinguistic data to shed light on the processing of coerced meanings. Let’s start with studies on complement coercion.

2.3.7 Studies on complement coercion


(72) a. The journalist began the article after his coffee break. (coercion)
b. The journalist wrote the article after his coffee break. (control)

Complement coercion such as began the article in (72a) was compared to a non-coercing control condition wrote the article in (72b). The latter used the verb which readers (in completion norms) most often ascribed to the eventive interpretation of the coerced expression. Reading times indicate that (72a) was more difficult to process than (72b) although plausibility judgments show that both were considered semantically well formed. In a word-by-word self-paced reading study, participants took longer to read article and after in (72a) than in (72b) (McElree, et al. 2001). In eye-tracking (Traxler, et al. 2002) and Pickering, et al. (2005) found reliable differences at the NP the article and the two words following it. Interestingly, these differences only showed up rather late in total reading times and not in first-pass. At least in syntax, late measures are usually taken to indicate processing difficulty having to do with reanalysis rather than the initial interpretation (Rayner, et al. 1989). It thus seems as if processing difficulty due to complement coercion reflects repairing a type mismatch rather than detecting it.

There is, however, a problem with the control condition in (72b). Note that the two sentences are not fully synonymous. (72a) means that the author began writing the article, but (72b) means that he finished it. Does the processing difficulty reflect differences in aspectual properties rather than complement coercion? Pickering, et al. (2005) used an additional control condition such as began writing the article and found no difference between this condition and the old control condition (72b), but coercion in (72a) was again slower than the control conditions. Another concern has to do with the different verbs preceding the target region. Egg (2005)’s analysis, for example, assumed ambiguity in eventive verbs; begin can be both a control verb and a NP selecting verb while write is a simple activity verb introducing one event. Is the apparent coercion effect due to the different verbs? Traxler, et al. (2002) compared conditions containing entity-denoting NP types like the article with event-denoting NP complements like the interview. In a two-by-two factorial design they crossed the factors NP-type and verb-type. The only condition that was slow was an entity-denoting NP after an event-selecting verb. This finding provides strong evidence that it is indeed the (re-) interpretation of the argument that causes difficulty.

16. Although Dölling (2003b) argues that coercion phenomena do not tell us anything about compositional versus non-compositional interpretation because they can be handled within strictly compositional semantic theories. In fact, many if not most of the coercion accounts that have been presented in the last section claim to fulfill strict compositionality.
Coercion effects generalize to other NP types and to languages other than English. In an eyetracking study, McElree, et al. (2006) found difficulty with proper nouns which are coerced into an event reading, *… began Dickens versus … met/read Dickens*. While the latter two conditions were read equally fast, the former caused difficulty. In an eye-tracking study on German Scheepers, et al.(2004) used materials like in (73).

(73) Der Student began das Buch mit großer Freude zu lesen.
The student began the book with great pleasure to read

They found a complement coercion effect at the complement even though the rest of the VP to read came later in the sentence.

To sum up, complement coercion reliably leads to disruptions during online comprehension. This difficulty appears to reflect rather late processes and has been found in different kinds of constructions, with different kinds of NP types and also in different languages. But what exactly makes complement coercion difficult? Let's look at studies that can inform us about the processes that underly complement coercion.

Underlying processes
What cognitive processes cause difficulty in complement coercion? Is it mismatch detection when trying to compose an entity-denoting NP with an event-selecting verb? Is it retrieving alternative reinterpretation candidates from conceptual knowledge? Is it selecting the most appropriate one? Or is it fitting a type-shift operator in the semantic representation? All these processes played important roles in semantic accounts of type coercion in the preceding section. And all of them are plausible candidates that could be responsible for the measured difficulty. We will look at the existing studies to decide between these alternatives.

The first option, mismatch detection, should be reflected by early measures. This doesn't fit very well with the delayed and sustained complement coercion effect that has been observed in the above mentioned reading time studies. What about the other three candidates? In an interesting eyetracking study, Traxler, et al. (2005) investigated what kind of discourse context eliminates the coercion effect. They placed the required activity in the preceding context assuming that this would eliminate the processing cost if the difficulty involved retrieving or inferring an appropriate activity. Consider the example in (75).

(75) a. The author began the book.
b. The author disliked the book.
c. The author wrote the book.

eyetracking study on German Scheepers, et al.(2004) used materials like in (73).

(73) Der Student began das Buch mit großer Freude zu lesen.
The student began the book with great pleasure to read

They found a complement coercion effect at the complement even though the rest of the VP to read came later in the sentence.

To sum up, complement coercion reliably leads to disruptions during online comprehension. This difficulty appears to reflect rather late processes and has been found in different kinds of constructions, with different kinds of NP types and also in different languages. But what exactly makes complement coercion difficult? Let's look at studies that can inform us about the processes that underly complement coercion.

Underlying processes
What cognitive processes cause difficulty in complement coercion? Is it mismatch detection when trying to compose an entity-denoting NP with an event-selecting verb? Is it retrieving alternative reinterpretation candidates from conceptual knowledge? Is it selecting the most appropriate one? Or is it fitting a type-shift operator in the semantic representation? All these processes played important roles in semantic accounts of type coercion in the preceding section. And all of them are plausible candidates that could be responsible for the measured difficulty. We will look at the existing studies to decide between these alternatives.

The first option, mismatch detection, should be reflected by early measures. This doesn't fit very well with the delayed and sustained complement coercion effect that has been observed in the above mentioned reading time studies. What about the other three candidates? In an interesting eyetracking study, Traxler, et al. (2005) investigated what kind of discourse context eliminates the coercion effect. They placed the required activity in the preceding context assuming that this would eliminate the processing cost if the difficulty involved retrieving or inferring an appropriate activity. Consider the example in (75).

(75) a. The author began the book.
b. The author disliked the book.
c. The author wrote the book.

They monitored brain activity while reading the complement and found a double dissociation of brain areas active in processing coercion and semantic anomaly.

Semantic anomaly yielded increased activity in the left temporal cortex corresponding to the M350. This is an MEG subcomponent of the N400 which is well known from ERP research (Kutas & Hillyard 1980). The N400 is a component which indexes difficulty in lexical semantic processing. Pykkänen & Marantz (2003) has hypothesized that the M350 is most likely a subcomponent of the N400 because these components are not always modulated by the same sort of experimental manipulations.

By contrast, coercion did not modulate brain activity in any language area, but elicited increased amplitudes in ventromedial prefrontal cortex. This new component which McElree and Pykkänen called the anterior midline field (AMF) is taken to reflect non-compositional sentence-level interpretation. Outside language, the neural
generators of the AMF are active in social cognition and in theory of mind (Rowe, et al. 2001). The latter plays an important role in false belief tasks. This is interesting because coercion in (76a) relies on inferring what authors typically do to books or to put it slightly differently, on finding out what their intentional stance is with respect to the direct object. Bringing this explanation together with the findings of Traxler, et al. (2005) we get a new perspective on the observed context effects. While accomplishments are clearly connected to the goals of the agent, activities are not: you may be building although you don’t even realize what you are actually doing. But you aren’t building a house unless you intend to do so. Given this observation it isn’t very surprising that mentioning the activity doesn’t eliminate the difficulty of constructing an accomplishment sense. The reader still has to infer what the agent intended to achieve by engaging in that activity.

Extending these findings, Pylkkänen, et al. (2009) investigated the generalizability of this effect. They varied the syntactic context of coercion by embedding the event-selecting verbs inside adjectives, as in (77a/b) and changed the task from an end of sentence acceptability judgment to offline comprehension questions.

(76) a. The ambitious artist envisioned the sculpture finishable in time for the exhibition.

b. The ambitious artist envisioned the carving finishable in time for the exhibition.

Again, they found a coercion effect similar to their previous study in location but about 250ms later. This shows that even without a decision task and in a different construction, the AMF effect is a stable neural correlate of complement coercion.

The results of the MEG studies can be compared with the findings of Choma (2006) and Kuperberg, et al. (2009) who conducted ERP experiments using the same design as in (76). In their studies, both coerced and anomalous sentences elicited a N400 which suggests that this type of coercion is resolved at the lexical level (for further discussion concerning the functional interpretation of the N400 see Chapter 7). An N400 effect is expected especially within the coercion account of Pustejovsky (1995) who assumes that complement coercion is computed using the lexical entry to come up with a new sense of the book. At first sight, these findings are in contrast to Pylkkänen & McElree (2007) who didn’t find a M350 in coercion. We have to be careful in interpreting this difference, however, because the M350 and the N400 are merely correlates.

2.3.8 Studies on aspectual coercion

In this section I will present an overview of the psycholinguistic literature on aspectual coercion. There are only few studies and the findings are rather mixed.

Pñango, et al. (1999)

Piñango, et al. (1999) examined cross-modal lexical decisions to unrelated words during the reading of sentences like (78a) and (78b) to test whether aspectual coercion imposes enhanced processing load. The asterisks indicate the probe position.

(77) a. The insect hopped effortlessly until *it reached the end of the garden.

b. The insect glided effortlessly until *it reached the end of the garden.

At the probe position, lexical decision times were longer for semelfactives like (78a) than for activities like (78b). The authors argue that the increased load reflects the online computation of an iterative reading which they take to be an enriched interpretation compared to compositional interpretation in the control condition.

There are, however, two problems with the study. First, two factors have been confounded. Besides the comparison between iterative and non-iterative readings the authors compare two different aspectual verb classes with each other. If activities are less difficult to process than semelfactives the same pattern of data is expected. Second, a closer look at the materials reveals a difference in plausibility between the two conditions. DeVelle (2004) collected plausibility ratings of the sentence materials. Her ratings show that across the board the coercion sentences were less plausible than the control sentences. This difference in plausibility is certainly another alternative explanation of the processing difficulty in Piñango, et al. (1999).

Todorova, et al. (2009a)

Todorova, et al. (2000a) pointed out some limitations of the Piñango, et al. (1999) study. They conducted an experiment employing incremental “stops make sense” judgments, in which participants evaluated whether the sentence continued to make sense as they paced themselves through it phrase by phrase. Besides bounded and unbounded eventualities they also tested two types of adverbial modifiers in a two-by-two design to rule out any effect that is due to the type of eventuality. The conditions are illustrated in (79a–d).

(78) a. Even though Howard sent a large check to his daughter for many years, she refused to accept his money.

b. Even though H. sent large checks to his daughter last year, …

c. Even though H. sent a large check to his daughter last year, …

d. Even though H. sent large checks to his daughter last year, …
The Processing of Events

Conditions (79a/c) have a singular direct object whereas (79b/d) have a plural object. Recall that the singular/plural manipulation changes the aspectual status from bounded to unbounded. The second manipulation was comparing durative (for many years) with locating adverbials (last year). If iterative coercion is taxing (79a) is expected to be more difficult than the other three conditions. This is exactly what Todorova et al. (2000a) found. Participants were more than twice as likely to reject sentences like (79a) than any of the control conditions (19% versus 7%, 8% and 9%). Furthermore, the associated reading times were reliably longer at the adverbial in (79a) than in (79b–d). The findings suggest that computing a coerced iterative reading does indeed increase processing load.

However, there are two caveats in order here. The first concerns the method used in this experiment. Stop-making sense judgments require a decision about the semantic status of a sentence after reading each phrase. With this procedure nothing can be left unspecified, that is all ambiguities have to be resolved. This is a highly unnatural situation and may thus be unlike ordinary reading. Second, the tested materials contained both semelfactive (e.g. kick) and achievement verbs (e.g. find). With the latter, durative adverbials are ungrammatical as illustrated in (80).

(79) a. “Even though I found my keys for ten minutes, …
b. Even though I found my keys last week, …

Thus, there is a possible confound in this study as well. It is not surprising that evaluating a sentence as semantically ill-formed yields higher rejection rates and longer reading times than evaluating sentences that are well formed. Unfortunately, we cannot conclude very much from the study.

Pickering, et al. (2006)

Pickering, et al. (2006) investigated whether aspectual coercion causes difficulty during ordinary reading without a concurrent task. They used the materials from Piñango, et al. (1999) and from a study by Todorova, et al. (2000b) and tested them each in a self-paced reading and in an eyetracking experiment. The Todorova, et al. (2000b) study was similar to Todorova, et al. (2000a) with the only difference that durative modifiers like for many years were replaced with quantificational phrases like every year. (a–d) is a sample item.

(80) a. Brian received a science book on his birthday every year …
b. Brian received science books on his birthday every year …
c. Brian received a science book on his birthday last year …
d. Brian received science books on his birthday last year …

Pickering et al. (2006) assume that in sentence (81a) coercion of an iterative reading is required. But this isn’t correct. Although the reading superficially looks like a coerced reading it can be derived in a strictly compositional fashion. Iteration is due to universal quantification over the achievement which doesn’t have to be reinterpreted at all. Consider (82b) which is a Neo-Davidsonian representation of (82a) with $e_2 \subseteq e_1$ expressing that event $e_2$ is part of event $e_1$.

\[
\begin{align*}
(81) & \quad a. \text{Brian received a science book every birthday.} \\
& \quad b. \forall e_1 . (\text{birthday}(e_1) \rightarrow \exists e_2 . (\text{receive}(e_2) \land e_2 \subseteq e_1 \land \text{recipient}(e_2, \text{brain}) \land \\
& \quad \exists y . (\text{Science book}(y) \land \text{theme}(e_2, y))))
\end{align*}
\]

For the above reason, since Experiments 3 and 4 didn’t test aspectual coercion they aren’t relevant for our purposes. For completeness, there were no differences in reading times between (81a–d) either in self-paced reading (Experiment 3) or in eye-tracking (Experiment 4).

In Experiments 1 and 2 there was also no indication of difficulty. Coercion was read as fast as the control condition. In analyzing the results, Pickering, et al. had to interpret null effects. This raises the question whether their study wasn’t sensitive enough to detect coercion effects. However, within the same experiments they included complement coercion sentences for which they obtained a complement coercion effect. This shows that the experiments were sensitive to reinterpretation costs. To explain the aspectual null-effects Pickering et al. claim that comprehenders routinely underspecify aspectual properties of an interpretation during normal reading. Only when reading is paired with a concurrent task readers commit to aspectual properties of an interpretation that they would otherwise leave unspecified.

This is, however, a rather strong claim given the fact that only one kind of aspectual coercion turning points into processes has been investigated experimentally so far. Just consider all the possible transitions in Figure 2.2! A much simpler alternative explanation is that coercion of a semelfactive into a process is particularly easy and therefore doesn’t cause a measurable disruption during ordinary reading. Another major point of critique concerns the materials that have been used. Piñango, et al.’s sentences contained verbs which are ambiguous between punctual and durative interpretations (e.g. hop, flash, yawn, nod, dive, …). Does to flash, for instance, denote a single flash? If some of these verbs have a process reading as well null-effects aren’t very surprising. All in all, aspectual properties have to be controlled before any far-reaching implications can be drawn from the experiments. This was done in the following study.

Brennan & Pylkkänen (2008a)

In a recent study, Brennan & Pylkkänen (2008a) investigated iterative readings of unambiguous semelfactives using self-paced reading and MEG. The materials were selected on the basis of a pretest which queried whether the candidate expressions
denoted a single event or multiple events on a scale from 1 (= single) to 7 (= multiple). The selected materials had a mean rating of 1.78 indicating they are clear instances of semelfactives. In their online experiments they compared coercion (83a) with control sentences (83b).

(82) 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.

In self-paced reading they found longer reading times on the region containing the verb in (83a) than in (83b). In the preceding region(s), however, the control condition was significantly slower than the coercion condition. In MEG, there were two neural correlates of the processing difficulty with different spatial and temporal profiles: first in right-lateral frontal and temporal sites between 350–400ms post stimulus, including regions involved in the processing of semantic anomaly. Second, in a later effect (450ms post onset) involving midline prefrontal regions that have previously been implicated in the resolution of complement coercion (Pylkkänen & McElree 2007). The authors interpret this two stage effect in the following way: in an initial step the verb is composed with the adverbial yielding a semantically ill-formed representation. The semantic mismatch is then pragmatically adjusted by iterative coercion. The former effect is indexed by the right lateral brain signature resembling a N400; the latter is reflected by the AMF, which indexes both semantic type shift (in complement coercion) and pragmatic type shift (in iterative coercion).

Unfortunately, the study also raises a number of questions. Concerning the materials, coercion and control differ with respect to tense. In the control condition the majority of items (19 out of 30) had an after x time adverbial. This type of adverbial locates the event time after the reference time which has to be introduced by an implicit event. Thus, two events are required in the control condition. The coercion condition requires only one, since reference time and event time coincide. At the adverbial region, control should thus be harder to comprehend than coercion. The reading time data corroborate this explanation. Up to and including the subject region, control was harder to comprehend than coercion. The reading time for coercion reveals a huge difference. The AMF due to complement coercion lasted for about 100ms and had a much bigger difference in amplitude between coercion and control than the aspectual AMF.

Finally, there is a methodological issue. In the reading time study no analysis of variance with items as random factor is reported. Item analyses are usually reported in the psycholinguistic literature in order to generalize an effect from single instances to the type of construction (Clark 1973). It would have been especially instructive to perform an item analysis to find out whether the different types of adverbials (after x time vs. at x time) in the control condition all show the same pattern.

A short summary
While processing difficulty in complement coercion is well-documented, the existing studies on aspectual coercion report rather mixed results. All of them focussed on only one type of aspectual coercion, namely iterative readings of punctual eventualities. It is not entirely clear whether this type of coercion increases processing load or not. Another desideratum is a systematic investigation of different coercion types. Similarly to research on complement coercion we should try to answer the following questions: what is the role of context information? What are the functional correlates of aspectual coercion in ERPs? In addition, a closer connection between psycholinguistic research and theoretical approaches is needed if we want to overcome the sceptical reflection of the existing findings by Dölling (2003b):

"[…] it is uncertain whether the processing effects can be viewed as resulting from a specific coercion operation performed to resolve an aspectual conflict between the verbal head and its adjunct. Instead, an interpretation of the results is conceivable such that they have their origin elsewhere.” (Dölling 2003b, p. 319)

2.4 Summary and conclusions
In this chapter we have traveled a long and winding road, so a brief recapitulation may be in order. Psychological investigations on event perception and cognition demonstrate that events are psychologically real. They have clear, perceptually identifiable boundaries and are mentally organized into hierarchical partonomic structures. Moreover, the developmental studies by Trabasso and Stein suggest that events are integrated into plan-goal structures and that planning provides the glue that connects isolated events into a coherent whole. Although the whole field is rather young with only a few studies yet, a growing interest in the psychology of events has to be attested.
In linguistics, a parallel development can be documented. Ever since Davidson (1967) events have become more and more fundamental in natural language semantics. Nowadays, they are commonly taken to be at the heart of the semantics of action sentences and they are taken to be ontological primitives together with processes and states. It is quite standard to distinguish five Aktionsarten: states, activities, accomplishments, achievements and semelfactives. These five aspectual classes are properties of complete sentences and are determined by the interplay of a lot of sentence internal material like verbs, their arguments and modifiers.

When items carrying aspectual information exhibit mismatching aspectual properties, they can be shifted into the required type. These type shifts are so called aspectual coercions. I distinguished between three types of semantic theories on aspectual reinterpretation. Operator-based theories assume that aspectual mismatch detection will lead to an update of the sentence representation with a type-shift operator that patches the apparent type-conflict. Underspecification accounts assume aspectually underspecified representations which are augmented with pragmatic information. The last type of theories, planning accounts, assume that aspectual interpretation has to do with integrating the events referred to in a discourse into a plan-goal structure. Within planning accounts, coercion can lead to drastic changes of the original plan-goal structure.

In presenting different types of aspectual semantic frameworks I have concentrated on EC by Hamm & van Lambalgen (2005). They provide a framework where the findings from cognitive psychology find a natural explanation. Hamm & van Lambalgen present a planning account which uses logic programming to compute temporal situation models. The derivations in the next chapter will exemplify how EC can be used as a processing theory from which we can derive predictions about aspectual coercion.

Finally, I have reviewed the psycholinguistic coercion literature. The main focus in the experimental studies has been on complement coercion, but aspectual coercion has recently also received some interest. With respect to aspectual coercion, however, the findings aren't conclusive.

In the rest of the book I will concentrate on the processing of aspectual coercion. The next chapter will introduce a general model of aspectual reanalysis. Based on this model, I will formulate alternative hypotheses. The hypotheses will be put to test in the subsequent chapters. In the final chapter, I will come back to the theoretical and psycholinguistic investigations that have been the topic of the present chapter and will discuss them in the light of the experimental results.

CHAPTER 3

Hypotheses and predictions

In this chapter I will lay out a general model of aspectual interpretation and discuss its processing consequences. The model covers both ordinary compositional interpretation where things work in a smooth fashion and the repair of local aspectual mismatch in case of aspectual coercion. I will discuss different possibilities that remain open in this model. Based on these, alternative hypotheses will be formulated which will be experimentally tested in the subsequent chapters. The hypotheses will be both about semantic and processing aspects of aspectual coercion and serve as the backbone of the monograph.

In Section 2 I will provide sample derivations of three different kinds of aspectual coercion. It will be illustrated that different kinds of coercion rest upon fairly different computations: non-monotonic updates of an eventuality in subtractive coercion, abstracting away from an eventuality in performing abstract type shifts and aspectual enrichment of an eventuality in additive coercion.

Finally, I have reviewed the psycholinguistic coercion literature. The main focus in the experimental studies has been on complement coercion, but aspectual coercion has recently also received some interest. With respect to aspectual coercion, however, the findings aren't conclusive.

3.1 A general model of aspectual interpretation

If semantic processing involves no type conflict, the semantic and the syntactic representation are strictly homomorphic to each other. This is depicted in the bottom part of Figure 3.1 which is termed "compositional analysis". Assuming incremental semantic interpretation, the syntactic and the semantic representation can be built in parallel: each new lexical item (carrying both syntactic and semantic information) is immediately attached to the phrase structure and composed with the existing semantic representation. Composition is used as a shorthand for the semantic counterpart (e.g. functional application) of the syntactic rules which are used to construct the syntactic representation. Even if the processor is recovering from a temporarily wrong attachment it still works within compositional analysis. This can be illustrated by a sentence that involves syntactic reanalysis like (1).

(1) The detective was watching a man with a revolver.
Since revolver is an implausible instrument of watching the PP with a revolver has to be attached to the NP a man. But since constructions such as (1) show a preference for attaching the PP to the VP the parse tree has to be reanalyzed from VP to NP attachment. Recovering from the garden-path not only involves syntactic reanalysis, but affects the meaning, too. Watching with an instrument has to be reanalyzed as watching with no instrument. The meaning of the syntactically reanalyzed sentence can, however, be determined on the sole basis of the meaning of the words and the way they are combined.

Figure 3.1. A model for aspectual interpretation and reinterpretation

Cases of aspectual reanalysis in which the composition of an incoming lexical item leads to a type conflict may introduce difficulty and can thus provide information about the processing up to this point. Type conflicts can be resolved in two general ways: either, the sentence-internal and perhaps even external context is semantically transformed to match the new input or the new input has to be adapted to match the sentence context. The former is the case in aspectual coercion, where the Aktionsart of the existing linguistic context is changed to match a new lexical aspectual item, while the latter is the case in complement coercion. In case of aspectual coercion, a number of coercion mechanisms are available (see Section 3.1.2) that allow shifting the meaning of the Aktionsart. In case repair was successful the argument is now of the right type and composition can take place.

The whole model is depicted in Figure 3.1. Nothing is presupposed about the time course and modularization of compositional analysis and aspectual reanalysis. For instance, reanalysis may even start before compositional analysis and mismatch detection are complete. This would be the case if the two steps were carried out in parallel allowing for partial output of the compositional analysis. Consider (4).

(4) John wrote [a letter] for hours.

Even if the processor hasn’t actually computed that John wrote a letter is an accomplishment it may output enhanced activation of an accomplishment analysis because it encountered a bounded (+SQA) direct object. Bringing the partial information together with the for-adverbial may therefore alert aspectual reanalysis to pick out the right kind of operator even though compositional analysis lacks some words behind. Similarly, the boxes are not intended to indicate any modular, but only functional differences between compositional analysis and aspectual reanalysis.

The model leaves open a number of questions concerning compositional analysis and aspectual reanalysis. Only the necessary ingredients are included that are relevant for explaining coercion. For instance, repair will only be carried out in case of a mismatch. I make the rather standard assumption that per default, semantic interpretation starts together with syntactic interpretation. In cases of aspectual reanalysis, semantic processing proceeds independently from attachment and doesn’t have an overt syntactic counterpart. The model serves as a starting point and will be refined throughout the book. In the following sections I will discuss the separate steps involved.

1. Strictly speaking, aspectual coercion could also be analyzed as a type-shift of the new input, but we intuitively talk about it in terms of context change whereas complement coercion is most naturally conceived of as reinterpreting the complement while the verbal meaning is kept fixed. For instance in a case of complement coercion like in (2) it is most natural to assume that the meaning of the book has to be coerced into an event reading to fit the required input format of begin:

(2) The author began the book.

(3) provides an example for aspectual coercion. This case is most naturally described in terms of changing the aspectual meaning from a punctual event to a process (= type shifting of the sentence context) when the coercion adverbial for an hour is encountered.

(3) The insect hopped for an hour.
3.1.1 What factors guide lexical aspectual access?

A number of lexical items and morphemes carry aspectual information. These are verbs (e.g., **build** vs. **potter about**), but also different kinds of adverbials (e.g., **in an hour** vs. **for an hour**), properties of arguments (e.g., **mass vs. count**), or aspectual markers (e.g., the **perfect**). Especially verbs are often to a high degree ambiguous between the different aspectual classes. Consider sentence (5):

(5) Hans baute ...
    Hans built ...

(5) already is a complete minimal sentence expressing an activity in the past. Interestingly, if we combine this sentence with the continuations in (6) the Aktionsart may change.

(6) a. ... an dem Haus herum.
    ... at the house around
    Hans did handicrafts about the house
b. ... ein Haus.
    ... a house
    Hans built a house
c. ... einen Unfall.
    ... an accident
    Hans caused an accident
d. ... zwei Jahre lang keinen Unfall.
    ... for two years no accident
    Hans caused no accident for two years

In (6a) **baute** is used as an activity because the **an NP herum-verb**-construction serves as an unambiguous cue. In (6b) it is an accomplishment; in (6c) we are dealing with an achievement usage of **baute** which is sort of idiomatic in German. Finally, in (6d) **baute** is used to express a state of not causing any accidents for a period of two years.

Given all these possible continuations what aspectual status should the parser attribute to the verb when it has not encountered the actual completion of the sentence yet? In principle, the aspectual status may either receive some aspectual specification right away or remain underspecified.

If an ambiguous input receives aspectual interpretation right away it has to be determined what factors guide the assignment of Aktionsart. Research on syntactic processing has shown that probabilistic information like frequency and discourse context plays an important role in determining the syntactic structure of a sentence (for an overview see e.g. Jurafsky (2003) and the references therein). Probabilistic parsing might be used in determining the Aktionsart as well. If we assume a probabilistic parser that predicts the right kind of continuation to go with a particular verb this continuation may be used to determine the aspectual properties of the sentence. For example, if the parser encounters the verb **build**, probabilistic information might be used to identify this verb as a good accomplishment verb, because **build** is most frequently used in accomplishment sentences. By contrast, if instead of **bauern** the verb **werkeln** (to do handicrafts) was used, due to frequency information of **werkeln** the parser should preferably assign it an activity reading. Especially under psycholinguistic accounts that assume no modular difference between syntactic and semantic parsing we would expect lexical frequency information to play a role in both domains. If aspectual parsing relies on probabilistic information we expect the **Probabilistic Parsing Hypothesis in (7)** to be correct:

(7) **Probabilistic Parsing Hypothesis (PPH)**

Aspectual properties of a lexical item are immediately determined on the basis of probabilistic information such as lexical aspectual frequency and context information.

A plausible alternative to this hypothesis consists in a parser which incrementally builds up a minimally updated aspectual representation whenever new lexical items come in. For instance, if the parser encounters a verb which is in principle ambiguous between an activity and an accomplishment (like **build**) it should introduce a process into the semantic representation because both activities and accomplishments include a preparatory process. If further downstream the sentence the processor encounters a direct object which can serve as incremental theme of an accomplishment (like a **house**) the semantic representation is updated by adding a culmination and a consequent state. By always choosing the minimal update of the semantic aspectual representation the danger of local semantic revision is minimized. Such a parser can be characterized by the **Lazy Parsing Hypothesis in (8)**:

(8) **Lazy Parsing Hypothesis (LPH)**

Every lexical item carrying aspectual information leads to a minimal update of the current Aktionsart such that it adds only as much to the model as is necessary to incorporate it into the semantic representation.

Assuming underspecification the processor does not decide on an Aktionsart when it encounters an item carrying aspectual information like the verb. Underspecification may be a viable option because it reduces the risk of aspectual misrepresentations. On the other hand, further semantic processing like the computation of a model is impossible and the lexical aspectual information has to be kept active in working memory. Contrary to the psycholinguistically common assumption that interpretation proceeds in an incremental fashion – that is on a word by word basis – underspecification implies that aspectual mismatch cannot be detected until aspectual properties of the sentence are determined. For any underspecification account whatsoever, we crucially need to know how big a portion of verb-argument-structure has to be present for aspectual
specification to be carried out. At the end of the sentence Aktionsart is certainly fully specified resulting in relatively clear intuitions about the temporal profiles of eventu-
talities. Therefore, underspecification has to be taken to be a local phenomenon. The *Aspectual Underspecification Hypothesis* is stated in (9):

9. **Aspectual Underspecification Hypothesis (AUH)**
   Semantic aspectual interpretation remains locally underspecified during normal comprehension.

Until proven false I will assume the PPH since this hypothesis fits best what is known about syntactic processing. Furthermore, it is psycholinguistically most interesting since it attributes a maximal degree of incrementality and predictive power to the parser.

The processing of aspectually ambiguous input will be the topic of Chapter 4. We will discuss the three theoretical alternatives by looking at the processing of aspectual ambiguity. This will further specify the bottom part of the model in Figure 3.1: the assignment of an Aktionsart to a new lexical item during compositional analysis.

### 3.1.2 What kinds of repair processes are there?

In this section I will lay out hypotheses concerning aspectual repair in the model (the reanalysis part of Figure 3.1). In the model repair is treated in its most general form as type transformation of the preceding sentence context. In the following, I will focus on different processes involved in type-shifting.

Starting with Moens & Steedman (1988) semantic theories on aspectual coercion always differentiated between a number of repair mechanisms for aspectual type mismatch. I will distinguish three kinds of coercion, partly adapting the classification proposed in Hamm & van Lambalgen (2005), Chapter 7.

In the event nucleus representation for Aktionsart introduced by Moens & Steedman (1988) there are some principled ways of transforming a nucleus into another. First, a whole eventuality might be used as the building block to construct a new one, without “touching” its subcomponents. For example, if an eventuality gets coerced into an iterative reading, this is done by recursively constructing an interval which consists of a number of iterations of the original eventuality as in (10).

10. \[ e \xrightarrow{\text{iterate}} e' \]

Interestingly, the coercing operation *iterate* can take any input type and shift it into a process. This is illustrated in (11) where *iterate* forced by a *for-adverbial* applies to an activity in (a), an accomplishment in (b), an achievement in (c) and a semelfactive in (d). It therefore provides a shifting device for purely *abstract type shifts*.

11. a. For half a year, John jogged in the park.
    b. For three weeks, John played the moonshine sonata.
    c. For many years, John won the race.
    d. For five minutes, John hiccupped.

Taking only a part of an eventuality, there are two possible modifications affecting the internal structure of the event nucleus: either the nucleus is updated with additional aspectual information (e.g. an *event* or a *process*) or only a subcomponent of the nucleus goes into the coerced meaning. These two kinds of coercion have been called *additive* and *subtractive coercion* by Hamm & van Lambalgen (2005).

12. John built for two years a house, until his credit canceled was.

13. John was building a house for two years, but then his credit failed.

The three coercion types can occur together in the same sentence as the famous example by Moens & Steedman (1988) in (14) illustrates. Intuitively, (14) is hard to comprehend. Is it because one kind of coercion is particularly difficult or because of multiple coercions piled on top of each other?

14. It took John more than three days to play the Minute Waltz in less than sixty seconds for more than an hour.

Example (14) involves abstract type shift since *for more than an hour* coerces play the *minute waltz* into an iterated reading by first transforming the accomplishment into a point and then iterating this point. The example also

---

2. States are irrelevant for our purposes here because they can in principle last arbitrarily long and in terms of coercion are hence unaffected by the *for-adverbial*. 
contains both subtractive and additive type shifts: Applying it took John two days to play the MW in less than 60s for more than 1 h involves two operations: first, turning the accomplishment into an achievement (= succeeding in playing the MW in less than 60s for more than 1h) is an instance of subtractive coercion. Second, this achievement has to be coerced back into an accomplishment by adding a preparatory phase (e.g. practicing) via additive coercion. Do all kinds of coercion mechanisms contribute the same amount of processing load?

A first possibility is that transformations of event nuclei induce difficulty regardless of whether there is any modification of the nucleus. Hence, in experiments studying the online processing of iterated readings we expect to find processing difficulty in coercing an event into an iterated reading. Of course, the same holds for other instances of abstract type shift like “slow motion” readings where an eventuality is stretched in time to become a process. This is stated in the Abstract Type Shifting Hypothesis in (15).

(15) Abstract Type Shifting Hypothesis (ATSH)

Abstract type shifting causes processing difficulty. Another potential source for difficulty in aspectual coercion could stem from removing parts of a complex event nucleus. Consider the lexical entry for an accomplishment VP like build a house. If the accomplishment is coerced into an activity, defeasible inferences which were valid before have to be canceled. As a result, the culminating event cannot be inferred anymore.

What are the processing consequences of withdrawing an inference? Does it induce difficulty? Note that defeasible reasoning under uncertainty is required in most everyday situations. Though rather exotic from the perspective of classical logic, it is quite common in terms of every-day reasoning and planning. Language use is characterized by an extensive use of pragmatic inferences to overcome the problem of insufficient information. The pragmatic literature following Grice (1967) is full of examples how interpretation makes excessive use of defeasible inferences (see e.g. Levinson (2002)).

In psycholinguistics there is a growing body of work studying the processing of implicatures and their cancelation (e.g. Noveck & Posada (2003); Bezuidenhout & Morris (2004); Bott & Noveck (2004); Storto & Tanenhaus (2005); Breheny, et al. (2006)). For instance, in sentences like some politicians are corrupt, in fact all are it has been investigated if canceling the generalized conversational implicature of some (= some, but not all) causes difficulty. Issues haven’t been settled yet and there is ongoing debate if generalized conversational implicatures are computed automatically (e.g. Storto & Tanenhaus (2005)) or if they only come in at later processing stages (e.g. Noveck & Posada (2003), Bott & Noveck (2004)). I don’t want to go into further details, but this line of study is interesting for the present purposes since it suggests that a non-monotonic reinterpretation of a lexical item like some (= some and possibly all) may enhance processing load.

In the case of subtractive aspectual coercion we are dealing with revisions at a supralexical level. The Non-Monotonic Update Hypothesis is stated in (16).

(16) Non-Monotonic Update Hypothesis (NMUH)

A non-monotonic revision of a discourse model causes processing difficulty. Finally, an additive update of a given event nucleus with processes, events or states may induce processing difficulty. For an additive update, contextual factors play a crucial role. In (17) a preparatory phase has to be inferred for the coerced accomplishment reading. The example illustrates that the contextual information of being at a party in (b) facilitates inferring the right kind of process (e.g. searching) compared to the example without such a context in (a).

(17) a. John spotted Mary in five minutes.
    b. At the crowded party, John spotted Mary in five minutes.

Additive coercion can be described in cognitive terms like the following: when a coercing phrase is encountered, the Aktionsart which is held active in working memory is updated with a slot for an additional eventuality. Since this piece of information is linguistically not realized it has to be filled in by looking for appropriate context information. In case the sentence is uttered out of context the comprehender has to rely on knowledge about a prototypical situation from declarative memory. The Aspectual Enrichment Hypothesis in (6) states that this kind of inferencing beyond the linguistic representation leads to processing difficulty.

(18) Aspectual Enrichment Hypothesis (AEH)

Enriching the aspectual representation with some additional eventuality causes processing difficulty.

We can further identify if processing difficulty is caused by filling in an empty slot of a specific aspectual type into the representation or by inferring the right kind of (concrete) eventuality from declarative memory. If the latter causes difficulty we would expect contextual influences to reduce processing difficulty in additive coercion. This question will be addressed in the context experiment in Chapter 4.

By investigating these different types of coercion it will be possible to further specify the upper part of the general model in Figure 3.1. Instead of just one general coercion operation three different kinds have been distinguished: abstract type shift, subtractive coercion and additive coercion. Whether each of them introduces processing difficulty will be investigated in three reading time studies in Chapter 5.

3.1.3 What is the processing domain for aspectual coercion?

Studying the processing of Aktionsart is particularly interesting, because there is a tension between assumptions that have been made in the semantic literature on Aktionsart
and in the processing literature. Studies on syntactic processing provide evidence that interpretation proceeds incrementally that is on a word by word basis. Does this hold for semantic processing at a supralexical level as well? (Rochri 1996, p. 251) formulated the principle of incrementality (the psycholinguistic perspective on syntactic processing) in the following way:

“The sentence processor operates in such a way as to maximize the interpretation and comprehension of the sentence at each stage of processing (i.e. as each lexical item is encountered).”

By contrast, in semantic theory Aktionsart is often treated as a property of whole VPs or even whole sentences. In Chapter 2 we have already encountered linguistic arguments why Aktionsart can only be determined at a phrasal level. I will call this the semantic perspective (Dowty 1979, p. 62):

“Not just verbs, but in fact whole verb phrases must be taken into account to distinguish activities from accomplishments. (In a certain sense, even whole sentences are involved...)”

If we extend Dowty’s claim to the processing of Aktionsart, we expect that for instance an unambiguous transitive verb on its own has no Aktionsart until it is composed with (at least) its internal argument. Consider the examples in (19). Erreichen (reach) is an unambiguous transitive achievement verb.

(19) a. Der Bergsteiger erreichte den Gipfel (*zwei Stunden lang). The mountaineer reached the top (*for two hours)
   b. Den Gipfel erreichte (*zwei Stunden lang) ein Schweizer Bergsteiger. The top reached (*for two hours) a Swiss mountaineer.
   c. Stolzer füllt erreichte (*zwei Stunden lang) ein Bergsteiger Full of pride reached (*for two hours) a mountaineer den Gipfel. the top

Note that all of the word order variants in (a), (b) and (c) are grammatical German sentences including the different positions the adverbial phrase can occupy as can be seen if we replace for two hours with a semantically appropriate adverbial like gestern Morgen (yesterday morning).

Because achievements are punctual they cannot be modified by a durative adverbial like zwei Stunden lang. Therefore, (19a) is ill-formed. The transitive verb has received all its arguments when the adverbial encountered, so both the semantic and the psycholinguistic perspective make the same prediction: the mismatch should be detected immediately at the adverbial phrase zwei Stunden lang.

In (19b), however, the adverbial intervenes between the verb plus its direct object (= the VP) and the subject. If indeed “whole sentences are involved” in determining the Aktionsart, then contra the psycholinguistic perspective, there should be no difficulty at the adverbial but only when the subject phrase is encountered. This is captured by the Late Aspectual Interpretation Hypothesis in (20). This hypothesis predicts processing difficulty due to aspectual mismatch and coercion only after the whole verb-argument structure is complete.

(20) Late Aspectual Interpretation Hypothesis (LAIH)

The smallest domain for computing Aktionsart consists in a verb with all its internal and external arguments.

If in (19b) mismatch detection takes place immediately at the adverbial there are still two possibilities as far as the processing domain of Aktionsart is concerned: the VP (the verb plus its direct object) or the achievement verb on its own. To decide between these two, consider sentence (19c). Here, the modifying adverbial intervenes between the verb and its arguments. If the semantic perspective is correct that Aktionsart is a property of a whole VP, in (19c) we expect no processing difficulty at the adverbial but only later in the sentence. This is stated in the Complete Verb Phrase Hypothesis in (21). Under (21) we expect to find processing difficulty introduced by aspectual mismatch and coercion only in case a whole verb phrase is present.

(21) Complete Verb Phrase Hypothesis (CVPH)

Only a verb with all its internal arguments is specified for Aktionsart.

If the adverbial can be semantically composed with the lexical verb we expect in (19c) immediate difficulty at the adverbial. Such a finding would be expected under an incremental perspective and is consistent with the Incremental Aspectual Interpretation Hypothesis in (22).

(22) Incremental Aspectual Interpretation Hypothesis (IAIH)

An aspectually unambiguous verb is the processing domain for Aktionsart.

3.1.4 How is semantic reanalysis carried out in the brain?

In this section different theoretical alternatives concerning the semantics of coercion will be discussed in the light of their assumed neurophysiological correlates. I will use event-related potentials (ERPs) to further elucidate the processing of aspectual mismatch and coercion. ERPs have very good temporal resolution and, moreover, may help us to determine the underlying processes of aspectual coercion.

In the semantic literature, aspectual coercion has been treated in very different ways. On the one hand, accounts like Rothstein (2004) treat aspectual coercion as
The Processing of Events

involving syntactic reanalysis by introducing additional structure (a type shifting operator) into the representation. A coerced sentence is therefore associated with a more complex syntactic structure than a non-coerced sentence. Coerced sentences could then be treated along the lines of garden-path phenomena in structural accounts like the Garden-Path Theory (Frazier 1978; Frazier 1987) and coercion would lead to processing difficulty which is functionally related to syntactic reanalysis. Such an account uniformly explains processing difficulty in the syntactic and in the semantic domain.

Purely semantic accounts of coercion, on the other hand, view aspectual coercion as affecting only the semantic representation. We can distinguish two families of such accounts.

Theories building upon underspecified representations treat aspectual coercion along the lines of lexical ambiguity. At certain positions in the syntactic representation (e.g. above the VP node) yet underspecified coercion operators are introduced. They are only determined as far as their input and output types are concerned. World knowledge and contextual information is needed to fully determine their actual semantic content or to choose the correct operator.

Another class of theories ground aspectual reanalysis in more general cognitive algorithms like planning. Recently, both Steedman (2002) and Hamm & van Lambalgen (2005) have proposed semantic analyses of tense and aspect which essentially link the semantics of temporality to the planning capacity in humans.

The various theoretical alternatives differ in terms of the cognitive processes that are involved. While syntacto-semantic accounts treat coercion as involving syntactic reanalysis, the two purely semantic options attribute coercion effects either to the lexicon or to more general inferencing.

Using EEG data we can test what functional processes underly aspectual coercion. After a very brief introduction to ERPs I will state three alternative hypotheses on the neurophysiological correlates of coercion (see Chapter 7 for a much more detailed introduction).

Evoked activity in the EEG: ERPs
Event related potentials are nowadays a fairly standard method to study the online processing of language. ERPs consist of that part of the EEG which is correlated with the onset of an internal or external stimulus. There are four ERP components that have been standardly related to language processing, each with its unique properties and distinct functional interpretation. The N400 component has been associated with lexical access (e.g. Kutas & Hillyard (1980); Kutas & Hillyard (1984); Kutas & van Petten (2006)) and is probably generated in the left temporal cortex (e.g. Halgren, et al. (2002)). The P600 has been attributed to syntactic integration difficulty (Osterhout & Holcomb 1992, Kaan, et al. 2000) and can serve as a marker of syntactic reanalysis (Friederici 2002). Brain imaging data suggest that the P600 is generated in the left frontal lobe in Brodmann areas 44 and 45 (Broca’s area). Furthermore, two negative components at left anterior electrode positions have also been claimed to reflect language processing. While an early LAN component has been connected to morphosyntactic processing (e.g. Friederici, et al. (1993)), the working memory LAN has typically been claimed to reflect working memory load while processing linguistic material (e.g. Münte, et al. (1998)).

Predictions based on the different coercion accounts
Coercion is triggered by the occurrence of a semantic mismatch. Thus, all theories predict an N400 effect, since the coercing stimulus is semantically unexpected compared to a non-coercing control. However, for the following step – aspectual semantic repair – the different theoretical options diverge on what kind of processes are hypothesized to play a role.

If aspectual coercion implies an update of the syntactic structure we expect to find neurophysiological correlates of syntactic reanalysis in aspectual coercion. This is captured by the Syntactic Reanalysis Hypothesis in (19). If this hypothesis is correct, coercion should lead to a P600 effect in the evoked potentials.

(23) Syntactic Reanalysis Hypothesis (SRH)
Aspectsal coercion involves syntactic reanalysis.

If underspecification is the right account, a coercion operator is already structurally present independently of aspectual coercion. If no type-mismatch occurs, the operator gets interpreted as the identity function (\(\lambda x.x\)). In case of type-mismatch the appropriate operator (e.g. iterate versus stretch) has to be chosen. Coercion is therefore explained in terms of lexical ambiguity resolution as it is stated in the Lexical Disambiguation Hypothesis in (21). The only predicted effect is an N400 indicating lexical access to the right kind of operator.

(24) Lexical Disambiguation Hypothesis (LDH)
Aspectsal coercion consists in the lexical disambiguation of an underspecified coercion operator.

Finally, the semantic accounts of coercion that make use of the human planning capacity predict that brain areas involved in planning tasks should also be active in aspectual coercion. The Planning Hypothesis is stated in (23).

(25) Planning Hypothesis (PH)
Aspectsal coercion depends on planning.

In case of additive coercion, for instance, an existing plan has to be updated with additional subgoals that have to be inferred via forward chaining (moving from the initial to the goal state) or backwards tracking (going backwards from the goal to the initial
state). This should enhance working memory demands in sentences that involve coercion compared to those that do not. Neurophysiologically, this should be reflected in a working memory LAN.

To sum up, ERPs allow us to identify the cognitive processes that underly the processing of sentences involving aspectual coercion. Using the ERP method we want to decide between different theoretical options that assume fundamentally different processes in repairing aspectual semantic mismatch. The neurophysiological correlates of coercion will be the topic of Chapter 7.

3.2 Semantic derivations for different kinds of coercion

In this section I will provide example derivations for the three kinds of coercion that were discussed in Section 3.1.2. The derivations are carried out in a combined discourse representation theory (DRT)/event calculus (EC) framework (for an overview over DRT see Appendix B and Kamp & Reyle (1993); for EC see Hamm & van Lambalgen (2005)). A combined framework which provides us with a computational semantics paired with a potentially incremental syntax-semantics interface is worked out in Hamm, et al. (2005) and Andrade-Lotero (2006). I do not claim that the parser actually computes constraint logic programs but will use the formalism to speculate about algorithmic differences between the different kinds of coercion.

One reason for using the DRT/EC approach is to see how a combination of the accounts could be turned into a semantic processing theory which allows us to derive predictions about when and how aspectual coercion is carried out during language processing. Another interesting feature of the framework consists in the hierarchical architecture of the theory which allows us to identify three levels during incremental interpretation, all of which can be affected by coercion. The three different kinds of representations are: (1) the incrementally constructed syntactic parse tree, (2) a discourse representation structure (henceforth DRS) of the syntactic tree and (3) a genuinely semantic representation of temporality in the EC which allows the computation of the temporal profile of a discourse (see appendix A Hamm & van Lambalgen (2003) for details of the formalism). This three step procedure is by and large adopted from Andrade-Lotero (2007) and Hamm et al. (2005).

Another reason is that a combination of DRT/EC offers a complete treatment of two essential aspects of semantic processing. DRT provides us with disambiguated representations of discourse (see Appendix B for a very brief introduction to DRT). This is needed for the resolution of scope ambiguities and of anaphorical relations in discourse processing. The way it is used here, DRT serves as the syntax-semantics interface much in the spirit of logical form of May (1985). EC, on the other hand, computes computationally tractable mental models and allows us to derive inferences from them. The example in (26) illustrates that both components are needed.

(26) Fritz rief oft am Montag an.
Fritz called often on Monday.

The sentence is scope ambiguous between a reading in which it is often the case that Fritz calls on Mondays (wide scope reading of often) and a reading in which last Monday Fritz was often calling (narrow scope reading of often). Before a model can be computed the sentence has to be disambiguated. Let’s assume that the wide scope of often is chosen. What kind of situation do readers construct given this reading? Most probably, they will construct a model in which there are extraordinary many phone calls by Fritz on Mondays compared to other days of the week. But that’s it, they will compute a model that is minimal. Unless there is evidence to the contrary, for instance, the reader will not assume that Fritz sang an aria every morning although this is fully consistent with the truth conditions of the sentence.

The two step procedure of first disambiguating the syntactic representation and then computing a model for it is, however, in conflict with recent psycholinguistic models which assume parallel construction of the syntactic and the semantic representation. But unfortunately up to date there is no formal semantic theory that allows us to model the construction process of the semantic representation independently of syntax and at the same time to capture semantic phenomena like scope, anaphora resolution and temporality in language. I will therefore stick to the two step procedure that allows word-by-word processing with broad semantic coverage.

In the following sections examples of subtractive coercion, abstract type shift and additive coercion will be derived. I will start with subtractive coercion because it will be argued that this kind of coercion can be resolved on purely lexical grounds. Abstract type shift makes use of fairly common semantic transformations while additive coercion is resolved by enriching the EC representation on the basis of extralinguistic information.

3.2.1 Subtractive coercion

Let’s derive the meaning of a simple sentence (27) involving subtractive coercion. The computation of the meaning will proceed phrase-by-phrase, from left to right, just like in ordinary language comprehension.

(27) John/ baute das Haus/ zwei Wochen lang/ bis der Kredit
John built the house for two weeks until the credit
gekündigt wurde.
was canceled.

In DRT, the starting point to give a temporal discourse a semantic representation is syntax. I will use the classic CP-IP analysis of German sentence structure (see von Stechow & Sternefeld (1988)). The construction rules modify the tree in such a way
that with each step parts of the syntactic representation get translated into a DRS until the complete tree has been translated into a complete DRS. A DRS is a box-like semantic representation and consists of two parts: (1) a universe, containing the discourse referents (objects in a first-order model universe like individual variables, variables for instants of time, etc.) and (2) a set of reducible and irreducible conditions. A DRS is complete when it contains only irreducible conditions (which are predicates in a first-order model).

When the parser encounters Johann it will immediately predict that Johann is the subject of the sentence (as numerous studies on subject-object-ambiguities in German show). Thus, the derivation starts with the syntactic tree in (28).

(28)

\[
\begin{array}{c}
\text{CP} \\
\text{NP} \\
\text{PN} \\
\text{Johann}_1 \\
\text{NP} \\
\text{VP} \\
\text{IP} \\
\end{array}
\]

I assume that every applicable construction rule immediately applies. Thus, the construction rule for proper nouns translates (28) into the DRS in (29).

(29)

\[
\begin{array}{c}
\text{x} \\
\text{John(x)} \\
\text{CP} \\
\text{x} \\
\text{IP} \\
\text{VP} \\
\end{array}
\]

The irreducible part of the DRS (namely \(\text{John(x)}\)) can be directly translated into an EC-representation, namely an integrity constraint (30) introducing a time-dependent property \(\text{John}(x)\), which can be unified with the constant John. Informally speaking, this integrity constraint (30) is an instruction to build a model containing the constant John. Because the model is minimal it will contain nothing else than John. It is noteworthy, however that this computation would have failed if the reader had assumed that no such person as John exists. The declarative memory of the reader would then have included a proposition like \(\forall x. \neg \text{John}(x)\) in contradiction with \(\exists x. \text{John}(x)\). Thus, EC elegantly allows us to model pragmatic influences in language comprehension.

(30) \(\text{HoldsAt(John(x), t), x = John}\) succeeds

When the reader encounters the next phrase (baute das Haus) the parser is already dealing with a complete sentence. The parse tree is given in (31).

(31)

\[
\begin{array}{c}
\text{CP} \\
\text{NP} \\
\text{PN} \\
\text{Johann}_1 \\
\text{NP} \\
\text{VP} \\
\text{IP} \\
\text{te} \\
\text{NP} \\
\text{V} \\
\text{V'} \\
\text{das Haus} \\
\text{t}_2 \\
\text{t}_1 \\
\end{array}
\]

Attaching the tensed verb plus its object introduces all facets of temporality. In DRT, these are captured by four features:

1. **TENSE** is a property of sentences and can have the values *past*, *present* or *future*. The main criterion to distinguish between different tenses is the relation between the time of the eventuality referred to by the sentence and the utterance time (\(n\) from now on). In example (27) \(\text{TENSE} = \text{past}\) can be read off the German morpheme "te". It is determined by the INFL node of the syntactic representation. The corresponding construction rule introduces a condition that the reference time is before \(n\): \(t < n\).

2. **PERF** is used to distinguish between those expressions that refer to consequent states (+PERF) and those that refer to the preparatory phase or the culmination (−PERF). The perfect, for instance, refers to the resultant state (and therefore is +PERF) while the simple past refers to both the preparatory phase and the culmination of an accomplishment verb *build a house*. In example (27), the feature value (−PERF) is determined by the tensed verb *baute*. In particular, the PERF value is a property of the head of the CP. The corresponding construction rule introduces an event which lies within the reference time: \(e \subseteq t\).

3. **AKTIONSART**, is not contained in classic DRT where a stativity feature STAT is used instead. Since STAT does not allow to distinguish accomplishments from achievements and semelfactives either states from activities it is too
coarse to serve for my purposes. Instead, AKTIONSART encodes the standard five Aktionsarten and is a feature of the verb plus all its arguments. AKTIONSART consists of a triple (e.g. ⟨+, +, +⟩) encoding the event nucleus of a phrase. The first value specifies whether there is a preparatory process, the second provides information whether there is a culminating event and the third specifies whether there is a resultant state. The triple is determined on the basis of the the three features ± SQA, ± ADDTO and ± DUR and can be read off Table 3.1 (see Section 2.2.4 for a discussion and justification). ADDTO and DUR are verbal features. The arguments all carry their respective SQA features which add to a global SQA value. Global SQA receives a plus value just in case all the arguments carry a positive SQA value. On the basis of a Davidsonian (Davidson 1967) event semantics eventsualities are treated as ontological primitives. These Aktionsarten are a property of at least VP or even higher hierarchical projections. Aktionsart ⟨+, +, +⟩ introduces an eventuality of type accomplishment into the DRS and yields the condition: ev(x, y, t, n): build(x, y).

4. TP is an abbreviation for temporal perspective which refers to the reference time of Reichenbach (1947). In (27) the temporal perspective point TP_{pt} is equal to n, therefore the relevant construction rule yields the condition: TP_{pt}: = n.

<table>
<thead>
<tr>
<th>Aktionsart</th>
<th>AKTIONSART</th>
<th>ADDTO</th>
<th>DUR</th>
<th>SQA</th>
</tr>
</thead>
<tbody>
<tr>
<td>activity</td>
<td>⟨+, −, −⟩</td>
<td>−</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>semelfactive</td>
<td>⟨+, −, −⟩</td>
<td>−</td>
<td>−</td>
<td>+</td>
</tr>
<tr>
<td>state</td>
<td>⟨+, −, +⟩</td>
<td>−</td>
<td>+</td>
<td>−</td>
</tr>
<tr>
<td>state</td>
<td>⟨+, −, +⟩</td>
<td>−</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>activity</td>
<td>⟨+, −, −⟩</td>
<td>+</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>achievement</td>
<td>⟨+, +, +⟩</td>
<td>+</td>
<td>+</td>
<td>−</td>
</tr>
<tr>
<td>accomplishment</td>
<td>⟨+, +, +⟩</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

Now, everything is ready to provide a complete DRS (32) for the parse tree in (31) stating that there was a building event of the house whose agent was John and this event lies completely in the past.

(32) \[ x, y, ev(x, y, t, n) \]

TP_{pt}: = n

\[ t < n \]

\[ ev(x, y, t, n) \]

\[ John(x) \]

\[ house(y) \]

\[ ev(x, y, t, n) \]

\[ build(x, y) \]

Again skipping technical details (see Andrade-Lotero, 2006), this can be translated into an integrity constraint in the EC as in (33). Additionally a scenario for the accomplishment VP has to be constructed encoding all relevant parts of the event nucleus (34). This is done by accessing the meaning of the accomplishment build-a house in the mental lexicon.

(33) \[ \text{Happens}(\text{build}_{build}(x, y), t), y = House, \]

\[ \text{Happens}(\text{start}_{build}(x, y), t), t < n \text{ succeeds} \]

(34) **lexical entry (scenario) for build-a house**

(1) \[ \text{Initially}(\text{house}(a)) \]

(2) \[ \text{Initiates}(\text{start}_{build}(x, y), t) \]

(3) \[ \text{Initiates}(\text{finish}, house(c), t) \]

(4) \[ \text{Terminates}(\text{finish}_{build}(x, y), t) \]

(5) \[ \text{Happens}(\text{build}_{build}(x, y), t) \]

\[ \text{Happens}(\text{finish})(\text{house}(c), t) \rightarrow \]

\[ \text{Happens}(\text{finish}) \]

(6) \[ \text{Releases}(\text{start}, house(y), t) \]

(7) \[ \text{Happens}(\text{build}(x, y), t) \rightarrow \text{Trajectory}(\text{build}_{act}(x, y), t) \]

\[ \text{house}(y + g(d), d) \]

**Baute das Haus** is an accomplishment in the perfective aspect which means that the whole situation is viewed as a single whole, without distinction of the various phases that make up that situation (Comrie 1976). In EC, perfective Aktionsarten are modeled by means of hierarchical planning which embodies a hierarchical organization of events. In (35) an event \( e_{\text{build}(x, y)} \) is defined which encompasses all the ingredients of the accomplishment, namely a process fluent build which denotes the preparation, a start event that initiates the preparation and a finish event. The latter is the culmination.

(35) \[ \text{build a house}_{(x, +, +), \text{in the perfective (via hierarchical planning)}} \]

\[ \text{Happens}(\text{start}_{\text{build}}(x, y), t) \land \text{Happens}(\text{finish}_{\text{build}}(x, y), t) \land r < s \leq t \rightarrow \text{Happens}(\text{build}_{\text{act}}(x, y), t) \]

The EC representation in (35) of a perfective accomplishment has been proposed by Hamm & van Lambalgen (2005) for English. All the parts that make up the accomplishment, that is the preparation and the culmination, are required to be true to make the perfective accomplishment true. However, in contrast to English German has no grammaticalized progressive. As a consequence, German accomplishments in the preterite are less restricted in their interpretation than English accomplishments in the simple past. Note the contrast between (36a) and (36b).

(36) a. "John built a house but he didn’t finish it."

b. "Johann baute ein Haus, aber er stellte es nicht fertig."

John built a house but he made it not finish.
While the English example in (36a) expresses a contradiction, the German example in (36b) sounds much more acceptable. It seems to be acceptable (at least to me) if one assumes that there was an event that prevented John from completing the house. So perhaps, German accomplishments in the preterite should receive an interpretation that is different from (35).

As a second theoretical alternative let's assume that the culmination in a German perfective accomplishment is only true in the default interpretation. If the context contains a terminating event stop that lies before finish the inference of the culmination will be canceled.\(^3\) Otherwise – if the context doesn't specify a stop event – negation as failure will make stop false, therefore finish has to be true.

(37) \textit{build a house} \textit{in the perfective (via hierarchical planning)}

The translation into the EC depends on which representation of the perfective accomplishment in (35) or (37) is chosen. Computing a model will involve two different computations with differential complexity. Either, the stop event introduced by the \textit{for}-adverbial is unified with the stop event variable in the formula of the perfective accomplishment in (37) changing the Aktionsart from accomplishment to activity. This computation involves the same number of steps as combining an accomplishment with an \textit{in}-adverbial and should therefore not pose any problems for the processor. An alternative computation is adding the information of the \textit{for}-adverbial to the perfective accomplishment representation of (35). This will lead to the computation of a local contradiction which has to be resolved by restating the integrity constraint. The second alternative predicts a local breakdown of interpretation which should be reflected by severe processing difficulty. Let's look at both theoretical alternatives in turn.

Derivation I: A stop-event cancels finish

In EC, the \textit{for}-adverbial introduces both the integrity constraint in (40) and the scenario in (41).

(39) \[ x, y, ev_{\text{f}}, ev_{\text{e}}, t, n, mt \]

\[
\begin{align*}
TPpt: & = n \\
\text{tp} & = n \\
ev_{\text{f}} & \subseteq t \\
ev_{\text{e}} & = t \\
\text{John}(x) \\
\text{house}(y) \\
\text{for}(ev_{2(\text{for},\ldots,\ldots,\ldots)}) \\
\text{two weeks}(mt) \\
ev_{2(\text{for},\ldots,\ldots,\ldots)} & \subseteq ev_{1(\text{for},\ldots,\ldots)} \\
\text{build}(x,y) \\
\end{align*}
\]

The definition of \textit{for} in (41) states that the adverbial is true if there is a process fluent \(f\) which is stopped after an amount of time \(mt. time(x)\) is a predicate that defines how we measure time (see Definition 41 in Hamm & van Lambalgen (2005, p. 178–179)). What happens when we combine this information with the accomplishment? We can now unify \(f\) with the building activity \textit{build} which is the preparatory process of the accomplishment. The integrity constraint of the accomplishment and the adverbial is provided in (42).

\(^3\) Strictly speaking, \textit{stop} needs to be defined itself. A stop event terminating a process fluent \(f\) can only happen \textit{at time} \(t\) when \(f\) holds at \(t\). That is, \textit{stop} cannot be unified with any arbitrary event but has to be \(f\)-relevant.
Resolution of this constraint uses clauses from the scenarios of the perfective accomplishment in (37) and for in (41). For introduces a stop event which can be unified with the stop event variable in the scenario of the accomplishment. Thus $\text{Happens} \left( \text{stop}_{\text{build}}, t' \right)$ is true marking the disjunction true. From this it follows that we cannot derive anything about the culminating event $\text{Happens} \left( \text{finish}_{\text{build}}, t \right)$ anymore. To see that we have to distinguish two cases: if stop happened before finish, that is if $t' < t$, closed world reasoning makes finish false. But if we assume that stop happened at the same time as finish, that is if $t' = t$, both stop and finish will be true. For pragmatic reasons, the first case is, however, more plausible in our example. Building a house typically requires more time than two weeks. As a result, the most plausible interpretation of the sentence is one where John was engaged in a building activity that lasted for two weeks and then got interrupted by some event which prevented him from completing the house.

Finally, when the processor encounters bis der Kredit gekündigt wurde a new event is added to the DRS which abuts on the building activity (due to the semantics of bis). This event can be unified with the variable stop in (37). Thus in the final interpretation of the sentence the credit failure stops the building activity.

Interestingly, the above derivation of subtractive coercion is no more difficult regarding computational complexity than the derivation of a non-coercing sentence with an in-adverbial as in (43).

(43) Johann baute das Haus in zwei Jahren.
John built the house in two years.

This becomes obvious when we consider the defining conditions of the in-adverbial stated in (44).

(44) \text{lexical entry (scenario) for an in-adverbial}

\begin{align*}
(1) & \quad \text{Happens} \left( \text{finish}, t \right) \land \text{HoldsAt} \left( \text{time} \left( \text{mt}, t \right) \right) \rightarrow \text{in} \left( e, mt \right)
\end{align*}

An in-adverbial is true in case the culmination happened and the time which elapses between start and finish corresponds to the amount of time mt. Combining the in-adverbial with the accomplishment will make finish true. Happens \left( \text{stop}, t' \right)$ is false due to negation as failure. The computation is thus very similar to the coercing case. The only difference is that the disjunction $\text{Happens} \left( \text{finish}, t \right) \lor \text{Happens} \left( \text{stop}, t' \right)$ is true because finish happened. As far as the processing of for- and in-adverbials is concerned, assuming that the derivation just given is cognitively real for and in should not differ in difficulty.

If this should turn out to be correct we can derive a very interesting prediction about crosslinguistic variation. In English, perfective accomplishments have a more strict interpretation than in German. While in German the culmination is only implicated in English it is entailed. In English, subtractive coercion cannot be computed along the lines of Derivation I and (if possible at all) should thus be considerably more difficult than subtractive coercion in German.

Note that Derivation I treats subtractive coercion as a kind of sense ambiguity. The accomplishment has a telic default interpretation but can also receive an aletic interpretation if the context so requires. The alternative derivation of subtractive coercion in the next section uses the more rigid representation of perfective accomplishments in (35) and henceforth requires a more complex computation.

Derivation II: Revision and recomputation

Instead of resolving subtractive coercion by shifting the lexical meaning of the accomplishment into an activity reading it can also be achieved by an alternative, albeit more complex computation. The integrity constraint that has to be proven true is repeated in (45).

(45) \begin{align*}
\text{?HoldsAt} \left( \text{John}(x), t \right), \quad x = \text{John}, \quad \text{HoldsAt} \left( \text{House}(y), t \right), \quad y = \text{House}, \\
\text{Happens} \left( \text{finish}_{\text{build}}, t \right), \quad t < n, \quad \text{for} \left( f, \text{mt} \right), \quad mt = 2 \text{weeks succeeds}
\end{align*}

The fluent $f$ of the for-adverbial has to be unified with a process. Among the defining conditions of the perfective accomplishment in (35) occurs a process, namely build.

Unification yields the following scenario (46) of the for-adverbial.

(46) \begin{align*}
\text{for unified with build a house and two weeks} \\
\text{Happens} \left( \text{stop}_{\text{build}}, t \right) \land \text{HoldsAt} \left( \text{time}_{\text{build}} \left( 2w \right), t \right) \rightarrow \text{For} \left( \text{build}, 2w \right)
\end{align*}

Trying to make the integrity constraint of for succeed leads to an update of the program adding a contextually given stop event to the model that terminated the building activity two weeks after it had started. Since world knowledge contains the fact that the building activity has to go on for more than two weeks before a house can be finished, Happens \left( \text{finish}_{\text{build}}, t \right)$ is false on pragmatic grounds. This yields a contradiction with Happens \left( e_{\text{build}x,y,t'} \right) - the head of the perfective accomplishment in (35) - which can only be made true if the finish event did happen. Because the integrity constraint in (45) expresses a contradiction it is unsatisfiable.

The only possibility to recover from this mismatch is to choose a different meaning of build a house which corresponds to an activity. (47) is a logic program which is derived from a building activity via hierarchical planning.

(47) \begin{align*}
\text{build a house, \ldots, in the perfective (via hierarchical planning)} \\
\text{Happens} \left( \text{start}_{\text{build}}, t \right) \land \text{HoldsAt} \left( \text{build}, t \right) \land s < t \rightarrow \\
\text{Happens} \left( e_{\text{build}x,y,t} \right)
\end{align*}
Restating the integrity constraint by exchanging the accomplishment \(a_{\text{build}}(x, t)\) with \(a_{\text{build}}(x, t) \rightarrow a_{\text{build}}(x, t)\) leads to a successful computation of a model. This model contains a start event which initiates a building activity. The building activity drives a partially changing house. After two weeks a stop event occurs that aborts the building activity. As a consequence, the house remains incomplete. Finally, unifying stop event with the credit failure event we get the final interpretation which is basically the same as in Derivation I.

Both derivations of sentence (27) are possible within the DRT/EC framework. It is an empirical question which alternative is correct. Although both derivations lead to a non-monotonic update of the situation model, the computations are quite different. While the first predicts that subtractive coercion is relatively easy the second states that in subtractive coercion the processor at first fails in computing a situation model and that only after reformulating the query eventually succeeds in interpreting the sentence. As a consequence, the latter computation should be reflected by severe processing difficulty.

Example (27) was chosen in a way that the building activity was too short to complete the house. What would happen if the time interval stated by the adverbial lies within the prototypical duration of the preparation, say two years? At least in principle, it would then be possible to unify the stop event with the culmination. The resulting meaning would then be identical to John built the house in two years. But since the speaker could have used the stronger in-adverbial to convey the same meaning, this interpretation is also ruled out on pragmatic grounds. However, given an appropriate context, this meaning can be enforced. Consider (48).

\[23\text{ John built the house two years long. Als er endlich fertig war, zog er mit seiner Familie ein. He moved with his family in.} \]

### 3.2.2 Abstract type shift

A different kind of coercion – abstract type shift – is exemplified in sentences (49a) and (49b). In (49a) a semelfactive sneeze is coerced into an iterative process while in (49b) an accomplishment load the wheelbarrow is transformed via abstract type shift.\(^4\)

\[49\]
\[
\begin{align*}
\text{a. } & \text{Johann nieste/ } \text{fünf Minuten lang. John sneezed for five minutes.} \\
\text{b. } & \text{Der Arbeiter belud die Schubkarre/ zwanzig Jahre lang. The worker loaded the wheelbarrow for twenty years.}
\end{align*}
\]

For (49b) to be true one needs to derive an activity predicate from a simple event. In EC there is a coding device which is explicitly dedicated to perform this type shift (event → fluent), namely imperfect nominalization (see Hamm & van Lambalgen (2005, Ch. 12)). Thus, in a sentence like (53) the gerund John's reaching the top of the Matterhorn has to be interpreted as a fluent and not as event.

\[53\]
\[
\begin{align*}
\text{a. } & \text{John's reaching the top of the Matterhorn was a surprise.} \\
\text{b. } & \text{John's reaching the top of the Matterhorn took place at ten.}
\end{align*}
\]

This is possible, although the gerund is derived via imperfect nominalization from a punctual achievement verb. In the analysis of abstract type-shifting we can employ
The Processing of Events

at the latest one hour after loading had started. As a result, at all times later than one
that the data base is the fluent (49b). What happens when it is combined with
ing the wheelbarrow whose agent was the worker and loading took between one min
the interval of five minutes lying completely in the past. It has further to be contextu
pragmatic constraint is included that loading a wheelbarrow takes at least one minute,
given in (57). Let’s assume that among the definitional statements of the scenario the
ment. The integrity constraint is stated in (56) and the constraint logic program is

(type shift. I will directly start with the EC representation of the perfective accomplish

The integrity constraint leads to the computation of a minimal model in which at least two sneezing events occurred: one at the beginning and another at the end of the interval of five minutes. It has further to be contextually specified how many events have to happen to be legitimized to speak of John’s five minute long sneezing. Thus, unlike subtractive coercion, abstract type shifting depends on world knowledge about typical situations in which the semelfactive holds. This becomes immediately clear if we compare sentence (52) with a sentence like (55) which is most compatible with an interpretation where only one sneezing event occurred which spanned an extraordinary long time.

In (49b), reanalysis proceeds by first coercing the accomplishment load the wheelbarrow into a semelfactive which is then turned into an activity by abstract type shift. I will directly start with the EC representation of the perfective accomplishment.

Computing an iterative reading II: Accomplishment \(\rightarrow\) activity

In (49b), reanalysis proceeds by first coercing the accomplishment the worker loaded the wheelbarrow into a semelfactive which is then turned into an activity by abstract type shift. I will directly start with the EC representation of the perfective accomplishment. The integrity constraint is stated in (56) and the constraint logic program is given in (57). Let’s assume that among the definitional statements of the scenario the pragmatic constraint is included that loading a wheelbarrow takes at least one minute, but no longer than one hour.

The integrity constraint succeeds if at some time in the past there was an event of loading the wheelbarrow whose agent was the worker and loading took between one minute and one hour. This is the minimal model corresponding to the first part of sentence (49b). What happens when it is combined with for twenty years?

The for-adverbial for(f, 20 years) refers to a process. The only process available in the data base is the fluent load, so f will be unified with it. The for-adverbial requires that load holds when the stop event happens, that is twenty years after the start of loading. On the other hand, the perfective accomplishment states that finish happened at the latest one hour after loading had started. As a result, at all times later than one
The additional conditions in (61) state that the duration of an accomplishment $ev_2$ that is generated on the basis of the finding event $ev_1$ was at most as long as an amount of time of five minutes.

When we try to compute a minimal model in the EC we need to reanalyze the scenario of find the key. (62) is the integrity constraint for the complete sentence and (63) is the scenario for the achievement find the key. This scenario simply states that if one finds something the state in the world changes from not having it to having it.

(62) \[ \text{Happens}(\text{Find}(x, y), t), x = \text{John}, \text{HoldsAt}(\text{Key}(y), t), y = \text{Key}, \text{HoldsAt}(\sim \text{Have}(x, y), t), \text{In}(\text{Find}(x, y), 5\text{minutes}), t < n \text{ succeeds}\]

(63) \[
\text{Happens} (\text{Find} (x, y), t) \land \text{HoldsAt} (\text{Have} (x, y), t) \land \text{Search} (x, \text{key}) \land \text{In} (\text{Search} (x, \text{key}), 5 \text{minutes}), t \land \neg \text{Find} (x, y), t
\]

Following the instruction stated in the integrity constraint $\text{In} (\text{Find}(x, y), 5\text{minutes})$ has to be made true. For this, the lexical entry of the in-adverbial in (64) is needed. The corresponding scenario states that an event happens in five minutes if there is some process (encoded by a fluent $f$) which is finished at some time $t$ and at $t$ the process has an amount of time of five minutes.

(64) \[
\text{Happens} (f\text{inished}, t) \land \text{HoldsAt} (f\text{inished} (= 5 \text{ minutes}), t) \land \text{In} (f\text{inished} (= 5 \text{ minutes}), t)
\]

Trying to fulfill the two premises in the body of the clause one needs a process which is not present in the scenario of the achievement. This process has to be inferred relying upon world knowledge. For find the key a good candidate would be search the key. But other activities are conceivable, too. For instance, asking your friends can also lead to finding something. Possible candidates are supposedly selected on the basis of the affordances of the direct object. This knowledge is stored in declarative memory and has the form of a perfective accomplishment rule like (65).

(65) \[
\text{a rule of finding via searching} \\
\text{Happens} (\text{start-search}, r) \land \text{HoldsAt} (\text{search}, s) \land \text{happens} (f\text{inish-search}, l) \land r \leq s < t \rightarrow \text{Happens} (f\text{ind}, t)
\]

In the new scenario (66) where find the key is reanalyzed as an accomplishment, the inferred information is highlighted in bold face.

(66) \[
\text{Happens} (\text{find the key}_{(x, +, +)}, t) \\
(1) \text{Initially} (\sim \text{Have}(x, \text{key})) \\
(2) \text{Initiates} (\text{start}, \text{Search}(x, \text{key}), t) \\
(3) \text{Initiates} (\text{finish}, \text{Have}(x, \text{key}), t) \\
(4) \text{Terminates} (\text{finish}, \text{Search}(x, \text{key}), t) \\
(5) \text{HoldsAt} (\text{Search}(x, \text{key}), t) \land \text{HoldsAt} (\text{Have}(x, \text{key}), t) \rightarrow \text{Happens} (\text{finish})
\]

Compared to a real accomplishment like build a house the derived accomplishment of find the key lacks an incremental theme which involves partial change. Another way to formalize the derived accomplishment would have been to use the already searched space as an incremental theme. However, searching may be quite unsystematic and the search space almost infinite. John could for instance have searched only a little part of the whole search space before finding his keys. Nevertheless the coerced accomplishment in (59) would be semantically appropriate. For this reason I conceptualized the derived achievement in terms of an accomplishment which has no incremental theme (contra Hamm & van Lambalgen (2005)).

To conclude, the derivation shows that additive coercion is fundamentally different from subtractive coercion and abstract type shift. In this type of coercion the resolution of mismatch heavily depends on world knowledge which is stored in declarative memory. As a result, additive coercion is the only form of coercion which cannot be resolved while staying within the language faculty.

The different kinds of computations that have been outlined in this section had fairly different algorithmic properties. They differed with respect to the involved processes and with respect to the number of computational steps needed to arrive at the coerced meaning. Building upon these algorithmic properties, different degrees of computational complexity are expected in different types of coercion. But since it is an open question whether aspectual coercion causes processing difficulty at all it is high time to leave computational semantics and turn to real life: aspectual processing in the lab.

3.3 Coercion at the offline/online-boundary: Experiment 1

I will now come to the experimental part. There has been some psycholinguistic work on aspectual coercion but up to now there has been no systematic investigation of different kinds of coercion. We know hardly anything about additive and subtractive cases of coercion in real time processing. Moreover, for the already investigated cases
of abstract type shift it is rather controversial if coercion causes processing difficulty at all (see Chapter 5 for existing studies on abstract type shift).

In the preceding two sections I have proposed the hypotheses and laid out predictions about the online processing of coercion. The semantic derivations for the three different types of coercion suggest that every type of coercion rests upon fairly different kinds of computational mechanisms. Thus, instead of treating aspeccual coercion as a uniform phenomenon I will compare the processing consequences of different kinds of coercion to shed light on the complexity of these computations.

All three kinds could eventually be resolved. They can thus be contrasted with global aspeccual mismatch which does not allow for aspeccual reanalysis. Looking at sentences involving coercion I have been assuming that they introduce processing difficulty but can be semantically repaired. They should eventually turn out as semantically well formed. The first experiment (Experiment 1) was conducted to find out (1) if there really is difficulty in aspeccual coercion, (2) if the amount of difficulty varies across different kinds of coercion and (3) if the aspeccual mismatch can be repaired.

The last point is particularly important because to be able to study aspeccual reanalysis comprehenders have to arrive at a coerced reading. In the online experiments that are described in the following chapters overwhelmingly isolated sentences were tested on naive subjects in a timed task. Under these circumstances, successful reanalysis may be particularly difficult. Temporary ambiguities can indeed be very hard to resolve so during normal reading they might not receive an interpretation at all but can only be understood when there is enough time to “contemplate” the sentence (an example from syntax is a garden-path construction like [the daughter of the king’s son was proud of himself]). Therefore, we have to make sure that comprehenders are able to perform coercion in the constructions under investigation.

For these reasons in Experiment 1 I want to find initial evidence concerning the grammatical status of additive and subtractive cases of coercion which haven’t been studied so far. Furthermore, I want to investigate whether in a timed task there is indication of intuitive difficulty due to additive and subtractive coercion. In the experiment constructions like the ones in (67) were used.

The climber climbed the mountain in two hours from the north
In two hours, the mountaineer climbed the mountain on the northface.

The climber climbed the mountain for two hours from the north
For two hours, the mountaineer was climbing the mountain on the northface.

The climber reached the top in two hours from the north
In two hours, the mountaineer reached the top from the northface.

Construction (67a) doesn’t require coercion. Here an unambiguous accomplishment verb is modified by an in-adverbial which needs an accomplishment. No type-mismatch occurs.

(67b) is a case of subtractive coercion. The accomplishment is modified by a for-adverbial which is atelic and therefore takes an activity as input. Thus, in order to integrate the adverbial the accomplishment has to be coerced into an activity. This is done by focussing solely on the preparatory process of the event nucleus (namely the climbing activity). That the accomplishment strips off its culmination and its consequent state can be seen when the sentence is continued with (68). This is a case of the so-called imperfective paradox (Dowty 1979). From (67b) it does not follow, that the climber actually reached the mountain top.

(68) ... als ihn ein Sturm zur Umkehr zwang.
... when a storm to go back forced.
... when a storm forced him to go back.

(67c) is a case of additive coercion. There is an unambiguous achievement verb erreichen (reach) which introduces a culmination (the climber reaches the top) and a consequent state (the climber now being on top). When the achievement is combined with an in-adverbial, it has to be coerced into an accomplishment. For this, new semantic structure (a preparatory phase) has to be added to the aspeccual representation. In our example world knowledge suggests that this might be a climbing activity. But, suppressing the information from the northface for a moment, he also could have reached the top using a helicopter.

Finally, (67d) is a case of unresolvable aspeccual mismatch. The for-adverbial needs a durative argument, but the achievement is punctual in nature. This leads to an unresolvable type mismatch. If repair were possible in this example it would consist of deleting the culmination and the consequent state from the achievement. At the same time, a process would have to be inferred, which is not semantically determined yet.

5. Besteigen (climb up) cannot be used intransitively and is generally bad in sentence frames that require an activity verb. For instance, one cannot say *Hans besteigt gerade (= At the moment, Hans is climbing) which is possible with activities like Hans joggt gerade (= At the moment, Hans is jogging).
Thus, the reader would be forced to construct a totally distinct verb- or VP-meaning instead of the one which is literally present in the sentence.6

3.3.1 Method

A timed method that is particularly suited to gather information both about the intuitive semantic status of a sentence and about the locus of difficulty is incremental makes sense judgements. In this task, the reader decides for each new phrase if the whole sentence fragment still is sensible. I will use these judgments as a starting point for the online experiments in the next chapters.

Materials

20 sets of four versions like in (67) were constructed, making sure that the time-interval stated in the adverbial was plausible as far as the duration of the process was concerned (a complete list of the items is provided in Appendix C). Slashes indicate segmentation. To make sure that only unambiguous achievement and accomplishment VPs were used in the experimental sentences the following criteria were applied: a VP was considered an achievement if the VP combined with a for-adverbial like in ‘reach the top for an hour’ should be unacceptable. By contrast, a VP counts as an accomplishment if it is subject to the imperfective paradox. That means that x climbed the mountain top follows from x climbed the mountain, but not from the VP in its progressive like in x was climbing the mountain.7 In addition, an accomplishment should allow modification by a

6. Interestingly, using the achievement in a progressive form like in (69) makes the sentence acceptable.

(69) Zwei Std. war der Kletterer dabei den Gipfel zu erreichen.

Two h. the climber was the top reaching
als ein Sturm aufkam.
when a storm emerged.

The climber was reaching the top for two hours when a storm emerged.

In (69) coercion is carried out in two separate steps: the progressive first turns the achievement into an (imperfective) accomplishment and then the for-adverbial coerces this accomplishment into an activity reading. In contrast to this two-step procedure, in the mismatch case the two steps would be happening at the same time. I will thus assume (following Egg 2005)) that only one coercion operation is possible at a given time. This is nevertheless consistent with many coercions being possible within one sentence.

7. To test this in German sentences lacking a morphologically realized progressive form, the sentences were put in a construction using x war dabei zu VPer (x was about to VP).
Participants
Twenty undergraduate students from Tübingen University (all native German speakers, 12 female, mean age = 22.3) participated in the experiment. They were naive to the purpose of the study. Each subject was paid 5 Euros for participation. The participants were randomly assigned to lists (five subjects per list).

Procedure
Experiment 1 used incremental stops making sense judgements for single sentences which were always followed by a yes/no-comprehension question. The sentences were displayed using the moving window technique (e.g. Haberlandt (1994)). With this technique the reader only sees the segment he is actually reading while all other other words of the sentence are masked (e.g. replaced by dashes). For each segment the reader had to choose between “yes, the sentence is still sensible” or “no, the sentence does not make sense anymore”. If the reader pressed the “no” button the trial was automatically aborted. If he pressed the “yes” button the next segment was displayed. Participants were instructed that they should make use of the “no” button only if they were sure that the sentence could not be continued in any sensible way. They were also told that they should read as fast as possible.

The experiment started with written instructions. These were followed by a practice block with ten sentences in which explicit feedback was provided. The practice items contained no aspectual violations. The experiment followed in a single block with a randomized order of sentences. After the experiment followed a five minute debriefing.

The experiment was conducted using a IBM compatible PC using E-Prime experimental software (www.pstnet.com). Responses were collected using a PST response box. The participants were tested individually in a quiet room. The experiment took approximately twenty five minutes.

Data analysis
Three dependent measures were analyzed: performance on the comprehension questions was analyzed to make sure that the participants read attentively. Rejection rates for each segment were used as an index of interpretability, while reaction times for the “yes” button presses served as indication of processing difficulty.

Reaction times for “yes” responses were corrected for outliers by trimming reaction times that were longer than 2500 ms by and times below 100 ms.

For statistical analysis repeated measures analyses of variance (ANOVAs) were computed with the within factors verb (accomplishment vs. achievement) and adverbial (for vs. in). Following Clark (1973) always two analyses were computed: using subjects and using items as the between factor ($F_1$ and $F_2$ respectively). Planned comparisons were computed using paired t-tests, again both by participants ($t_1$) and by items ($t_2$).

To reduce the danger of type 1 errors, these planned comparisons were computed under a Bonferroni corrected $\alpha$.

3.3.2 Results
Comprehension questions were answered correctly in 92% of all experimental trials (with little difference between conditions: mismatch 100%, additive coercion (ac) 87%, subtractive coercion (sc) 94% and control 90%). Thus, subjects read attentively. The nonsensical distractors lead to the abortion of the trial in 74% of all cases, whereas the sensible fillers were read to the end in 91%. This indicates that participants were computing interpretations while reading the sentences.

Cumulated rejection rates are depicted in Figure 3.2. As expected, up to the adverbial the sentences were accepted equally often across conditions. At the fourth segment (containing the critical adverbial) 46% of the trials were rejected in the aspectual mismatch condition, 11% in case of the additive coercion condition, 10% in the subtractive coercion condition and only 4% in case of no aspectual mismatch in the control condition.

Figure 3.2. Cumulated rejection rates in Exp. 1 (with 95% confidence intervals)
In this region, ANOVAs revealed a significant interaction between verb and adverbial \( F(1, 19) = 29.64; p < .01 \) and of adverbial \( F(1, 19) = 48.47; p < .01 \) and of adverbial \( F(1, 19) = 27.21; p < .01 \) and of adverbial \( F(1, 19) = 35.53; p < .01 \) were significant. Planned comparisons revealed a reliable difference between the mismatch and the additive coercion \( t(19) = 6.72; p < .01 \) and between the mismatch and the subtractive coercion condition \( t(19) = 7.60; p < .01 \). There was no statistically reliable difference between the additive coercion and the control condition \( t(19) = 1.93; p = .07 \).

On later segments, no statistical analyses were computed because of the different number of aborted trials across conditions. Therefore, only the numerical values will be reported. Until the end of the sentence the mismatch condition was rejected in 79% of all cases, roughly like the nonsensical filler sentences (74%). The coercion conditions were intermediate (38% for ac and 43% for sc respectively), while the control condition was rejected in only 9% of all cases (as often as the semantically well formed distractors).

Turning to the reaction times (RTs) for “yes, makes sense” judgments, RTs on the adverbial for the mismatch and the coercion conditions were longer than in the control condition. Diagram 3.3 depicts the RTs for the yes-answers in the four conditions.

![Diagram 3.3. Reading times for “yes, makes sense”-responses (corrected for outliers) in ms in Exp. 1 (with 95% confidence intervals)](image)

In the control condition the “yes” judgements had a mean RT of 1108 ms, while the two coercion conditions had mean RTs of 1499 ms in the additive and 1666 ms in the subtractive coercion condition. The mismatch condition also had a rather long mean RT (1554ms). Statistically, the interaction between verb and adverbial was marginally significant \( F(1, 19) = 3.07; p = .10 \) and there was a significant main effect of adverbial \( F(1, 19) = 8.79; p < .05 \) and of adverbial \( F(1, 19) = 7.42; p < .05 \). There was no indication of a main effect of verb \( F(1, 19) = 1.89; p = .19 \) and of verb \( F(1, 19) = 1.06; p = .32 \). Planned comparisons revealed a significant difference between the additive coercion and the control condition \( t(19) = 3.14; p < .025 \) and there was no reliable difference between the additive and the subtractive coercion condition \( t(19) < 1 \).

The RT effect was restricted to the segment containing the adverbial. On the following preposition there were hardly any numerical differences between conditions (values ranging between 728 and 785ms).

### 3.3.3 Discussion

Experiment 1 investigated the semantic acceptability of sentences involving local or global aspectual mismatch. Two kinds of coercion were compared. While additive coercion requires additional aspectual information, cases of subtractive coercion consist in focusing on just a part of the eventuality which is provided by the linguistic context.

Sentences containing an aspectual mismatch were overwhelmingly rejected as not sensible. By contrast, the subtractive and additive coercion cases were for the most part judged to be sensible. This shows that readers were able to perform coercion relatively fast. Nonetheless, higher rejection rates than in the control condition indicate that comprehenders had difficulty with coercion. Above that, reaction times indicated that both subtractive and additive coercion introduced difficulty right at the coercing adverbial. Although readers “stumbled” at the adverbial in the coercion conditions the sentences were majorally still judged to allow for a meaningful continuation.

One might ask, however, whether 60% acceptance rate is still enough to legitimately make the claim that the aspectual coercion sentences were in fact semantically acceptable. How can we account for these strikingly high rejection rates? First, we have to be aware that comprehenders in this experiment (but also those to follow) were reading decontextualized sentences. I think this makes coercion much less acceptable than in cases where the context actually supports the coerced meaning. Adding just a little bit of context, as it was done in the ERP study (Exp. 9) lead to 75% acceptance of additive coercion sentences. Second, self-paced reading isn’t exactly the most natural form of reading because readers do not have the possibility to regress.
their eyes to earlier parts of the sentence. Probably, in the coercion conditions participants may thus have got stuck and couldn't get at a sensible interpretation.

The results are interesting because existing studies on aspectual coercion provide a quite controversial picture concerning the question if there is enhanced processing load in aspectual coercion at all. While some studies found difficulty (e.g. Piñango et al. (1999); Todorova et al. (2000a), Husband, et al. (2006)) others did not (Pickering et al. 2006). Experiment 1 provides further evidence that there is difficulty in aspectual coercion. Moreover, the existing studies focussed solely on difficulty which is caused by the computation of iterative readings. Experiment 1 investigated two other types of coercion, subtractive and additive cases of aspectual semantic reanalysis. This enlarges the picture considerably and generalizes from a particular instance of coercion, namely iteration, to the broader claim that aspectual coercion causes difficulty in general.

Note, however, Experiment 1 did not investigate ordinary comprehension because the decision component introduces an additional task which is not present in ordinary language comprehension. The decisions may even lead to a fundamental change in language processing compared to ordinary reading. As each new segment is encountered the reader is forced to semantically combine the new input with the sentence context while in normal reading semantic composition might be delayed. For instance, assuming underspecification of Aktionsart for a moment, incremental make sense judgments would indicate fully specified Aktionsart on an incremental basis although this isn't the case in ordinary comprehension. The findings must therefore be interpreted rather carefully. Keeping these caveats in mind, the results of Experiment 1 suggest that cases of additive and subtractive coercion are very promising constructions to be used in experiments which investigate the resolution of aspectual mismatch in natural reading.

### 3.4 Summary and conclusion

In this chapter, I proposed a general processing model for aspectual semantic reanalysis. Within this model, I raised four questions which will be addressed in the remainder of this book.

The first question is how the processor deals with aspectually ambiguous information. This question is particularly interesting because many (if not most) verbs can be used to express nearly any Aktionsart. Three different hypotheses have been proposed which will be put to the test in the next chapter: the Aspctual Underspecification Hypothesis, the Lazy Parsing Hypothesis and the Probabilistic Parsing Hypothesis. The Probabilistic Parsing Hypothesis assumes a highly predictive parser as it has been proposed for syntax.

Second, I distinguished between different processes that might cause processing difficulty in aspectual coercion. Some kinds of semantic operations on an Aktionsart, like iteration have already been discussed in the psycholinguistic literature, while two other kinds of coercion, additive and subtractive coercion have not been systematically studied. Again, three hypotheses have been formulated which will be the topic of Chapter 5: the Abstract Type Shifting Hypothesis which attributes difficulty to abstract type shift, the Non-Monotonic Update Hypothesis which deals with subtractive coercion and the Aspctual Enrichment Hypothesis which assumes difficulty due to additive coercion. All three are plausible candidates for processing difficulty in coercion.

Third, it is an open question if compositional semantic interpretation which is needed for aspectual coercion works in an incremental fashion like attachment. How big a portion of verb-argument structure is needed for Aktionsart lead to three alternative hypotheses and will be the topic of Chapter 6: the Incremental Aspctual Interpretation Hypothesis, the Complete Verb Phrase Hypothesis and the Late Aspctual Interpretation Hypothesis. For the time being, I will entertain the Incremental Aspctual Interpretation Hypothesis since it is in accordance with the incrementality assumption from syntax.

Finally, an EEG study will get access to the neurophysiology of semantic repair in aspectual coercion. This last point is especially interesting because different theoretical options heavily rely on functionally different processes to explain how aspectual semantic reanalysis is carried out. Again, three hypotheses have been proposed: The first option consists in the Syntactic Reanalysis Hypothesis, the second in the Lexical Disambiguation Hypothesis and finally the Planning Hypothesis. These will be the topic of Chapter 7.

I derived aspectual reanalyses of three different kinds of coercion in a combined DRT/EC framework. This was done to demonstrate that there is a model which allows us to formally derive the coerced meanings in an incremental fashion – that is, as the parser encounters each new word or phrase. We found differences between abstract type shifting and subtractive coercion on the one hand and additive coercion on the other: while the former types of coercion can be solely resolved on the basis of information contained in the linguistic representation, additive coercion crucially involves the update of the scenario with contextual information. Furthermore, for subtractive coercion two alternative ways were identified how reanalysis could be achieved. These alternatives differed in terms of computational complexity. I will come back to these issues in Chapter 5.

Finally, using incremental makes sense judgements it was established that the constructions that will be used are well suited for ordinary reading experiments. Additive and subtractive cases of coercion were judged sensible although reaction times indicated difficulty. Achievements modified by a for-adverbal provide a construction involving aspectual semantic mismatch which can be used to compare cases of local and global aspectual semantic mismatch.
CHAPTER 4

The access to lexical aspectual information

In this chapter I will investigate how the parser specifies the aspectual properties of new input in case of aspectual ambiguity (e.g. of a verb like build in build a house vs. build houses). To this end, two factors, frequency information and discourse context, will be studied which have been shown to affect syntactic disambiguation in garden path sentences. It will turn out that both of these factors have an influence on the Aktionsart assigned to ambiguous material, too. If a verb is predominantly used as an activity (like jog) it is initially assigned an activity reading, whereas accomplishment verbs (like build) are given an accomplishment reading even before they are combined with the obligatory direct object (Experiment 2). Contextual influences can override these preferences (Experiment 3). If contextual information comprises an event which can serve as culmination, activity verbs are initially taken as accomplishments. Taken together, this indicates that the factors that are relevant for the resolution of syntactic ambiguity also have an influence in semantic processing.

4.1 Ambiguity resolution in the aspectual domain

One of the core problems in psycholinguistic research concerns the question how the processor deals with ambiguous input given the vast amount of local and global ambiguity in natural language. In fact, ambiguities can arise at each level of linguistic representation. This is shown in the examples in (1).

(1) a. The bank of Scotland is near the river bank.
   b. Old men and women entered the room.
   c. One continent is served by every airline.
   d. Can you make pancakes?

I will have a look at each of these examples and discuss the linguistic representation at which ambiguity arises. This will allow us to decide to which category ambiguities in Aktionsart belong.¹

¹ There are also phonological ambiguities where a given phonological sequence can be interpreted in more than one way. (2) from Jackendoff (2002, p. 202) is an example.

2nd proofs
lexicon: Sentence (1a) contains a lexically ambiguous word (bank) which has two different meanings. Besides homonyms like bank there are also polysemous words like newspaper which have several senses like physical object versus institution. Homonymy and polysemy are very common in natural language: the 500 most frequent English content words for instance have a mean of 35 different senses according to the Oxford English Dictionary.

syntax: Example (1b) allows two different syntactic analyses. These are illustrated by the following two bracketings of (1c): […] old [men and women]] versus […] [old men] and [women]].

semantics: Sentence (1c) contains a scope ambiguity and can either mean (a) that there is one continent (say Europe) which is served by every airline or (b) that for every continent there is a potentially different airline that serves it. Note that this kind of ambiguity cannot arise at a lexical level because all the lexical items are unequivocal. Neither is the sentence syntactically ambiguous. Hence scope ambiguity arises at the phrasal semantic level. The two scope readings are represented by the following predicate logical translations: \( \exists x (\text{continent}(x) \land \forall y (\text{airline}(y) \rightarrow \text{serve}(y, x))) \) versus \( \forall y (\text{airline}(y) \rightarrow \exists x (\text{continent}(x) \land \text{serve}(y, x))) \).

pragmatics: Example (1d) comprises an ambiguity at the pragmatic level. The literal meaning consists in a question (= is the addressee able to make pancakes?) while the intended meaning may be a request.

Looking at sentences with a local aspectual ambiguity like (3) we have to ask at which linguistic level this particular kind of ambiguity arises.

(3) a. Hans baute an dem Haus herum.
Hans built at the house around
Hans pottered about the house

b. Hans baute ein Haus.
Hans built a house
Hans built a house

c. Hans baute einen Unfall.
Hans built an accident
Hans caused an accident

d. Hans baute die letzten zwei Jahre keinen Unfall.
Hans built for the last two years no accident
Hans caused no accident during the last two years

2. However, if the lexicon is assumed to include whole constructions (e.g., Goldberg (1995), (2005); Seidenberg & MacDonald (1999); Tomasello (2003)) one can in principle employ a lexicalist view on ambiguity resolution in the aspectual domain, too. But, as a starting point, I will stick to the assumption that the lexicon contains no lexical entries for whole constructions.

3. Pragmatic ambiguity is no option either because Aktionsart behaves fairly different from examples like (1e). That is, while in Example (1e) at the same time there is a literal meaning and a pragmatically derived meaning – word and speaker meaning in the terminology of Grice (1967) – in the aspectual cases in (3) the sentences are used in their literal meaning.
readings of multiply quantified sentences (Kurtzman & MacDonald 1993; Filik, et al. 2004). Since scope ambiguities are genuinely semantic in nature, it is worth exploring whether multiple constraints are at work in aspectual processing, too.

4.2 Factors that play a role in syntax

The predominant view on syntactic processing is the constraint-based framework which holds that use of language exploits multiple probabilistic constraints over various types of linguistic and non-linguistic information. I will focus on two factors that have been playing an especially prominent role within the psycholinguistic literature: lexical frequency and context information.

4.2.1 Frequency information in syntactic disambiguation

Frequencies of linguistic structure, especially linguistic structure related to lexical items have proved to play an important role in the disambiguation of structural ambiguities (for an overview see Jurafsky (2003)). For example, the verbs understand and suspect are both subcategorized for either a direct object noun phrase or a sentential complement, as in (5).

(5) a. John understands the child.
   b. John understands that Mary will not come and visit him.
   c. John suspects the child.
   d. John suspects that Mary will not come and visit him.

Although both verbs allow both subcategorization frames, they do so with different frequencies. While understand appears more often with a noun phrase argument, suspect is more frequently used with a sentential complement. Thus, relying upon lexical frequency information of the verb the beginning of the complement clause that Mary in (5b) is more difficult to comprehend than the noun phrase the child in (a), but the pattern is exactly reversed in (5c) versus (d).

Subcategorization frequencies for verbs can be computed from syntactically annotated corpora (Merlo 1994; Roland & Jurafsky 1998) or they can be determined by asking subjects to write sentences using the verbs (Connine, et al. 1984; Garnsey, et al. 1997). The frequencies provide us with an estimation of the conditional probabilities of the subcategorization frame given the verb P(frame|verb). These probabilities have repeatedly been shown to play a role in parsing: the higher the conditional probability of the frame, the more it will be preferred in disambiguation (e.g. Trueswell, et al. (1993), MacDonald, et al. (1994)). Influences of lexical frequency information in parsing has been shown with a variety of methods like self-paced reading (e.g. Trueswell (1996)), eye tracking in reading (e.g. Boland & Blodgett (2001), but see Pickering, et al. (2000)

4.2.2 Context information in syntactic disambiguation

Another factor that plays a prominent role in syntactic disambiguation is context information. Since the seminal work of Altmann, Crain and Steedman (Crain & Steedman 1985; Altmann & Steedman 1988) there has been a growing body of evidence that the right (or wrong) kind of context can eliminate (or induce) garden path effects. The referential theory states that it is the presuppositional affordances of competing syntactic analyses which leads the parser to choose one particular structure. The referential theory makes the strong prediction that discourse can shift initial preferences.

This can be illustrated with the example in (6).

(6) a. A boy and a girl were standing in front of the school building waiting for the bus. When a teacher came along he started talking to the boy.
   b. Two boys were standing in front of the school building waiting for the bus. When a teacher came along he started talking to one of the boys.

Let’s have a look at the target sentence without context first. This sentence is locally ambiguous between a preferred main clause and a dispreferred reduced relative clause reading. When the boy talked to by the teacher was in his class.

In (6) the garden path sentence is already embedded into two contexts (6a) and (6b) which differ in terms of referents introduced into the discourse model. Here comes the referential theory of Crain, Altmann and Steedman into play: while context (6a) only introduces one boy; in (6b) there are two. The use of a complex NP like the boy talked to by the teacher carries the presupposition that there are at least two boys, while a simple NP like the boy only presupposes the existence of a unique boy. Therefore, in terms of presuppositions, without context or with context (6a) the reduced relative interpretation carries more unsupported presuppositions than the main clause interpretation and is therefore disregarded. By contrast, with context (6b) the main clause interpretation becomes somewhat odd because a simple NP like the boy can’t single out a unique discourse referent given a discourse model which contains two boys, thus violating referential requirements. The referential theory therefore predicts for a different finding) and in the visual world paradigm (e.g. Snedeker & Trueswell (2004)). Interestingly, lexical frequency information is already used for syntactic disambiguation in preschool children (for an overview see Trueswell & Gleitman (2004)). Taken together, this shows that frequency information plays a key role in determining the preferred structure in cases of syntactic ambiguity.
Applying frequency and context information to aspectual ambiguity

The Processing of Events

The syntactic parser makes use of probabilistic information, but semantic processing both proceed incrementally in a word-by-word fashion, but that the factors in syntactic ambiguity resolution are different from those in aspectual semantics. The syntactic parser makes use of probabilistic information, but semantic processing differs in their assumptions about the relationship between ambiguity resolution in syntax and aspectual semantics. These hypotheses (which have already been briefly described in Chapter 3) differ in their assumptions about the relationship between ambiguity resolution in syntax and aspectual semantics. Therefore, this hypothesis implies distinct time-courses in syntactic and in semantic processing. While syntactic processing works in an incremental fashion, aspectual semantic processing is carried out at the end of the sentence. Furthermore, aspectual semantic processing is fundamentally different from syntactic parsing in terms of the factors which play a role: while in syntax several probabilistic factors are exploited, aspectual semantics does not need to employ probabilistic information like frequency information. These kinds of information aren't needed, because the aspectual semantic parser is not forced to predict anything on the basis of incomplete information, but always has all parts contributing to the meaning of the whole sentence available. Thus, underspecification assumes the syntactic and the aspectual semantic processor to be different from each other both in terms of time course and of factors affecting disambiguation.

4.3 Applying frequency and context information to aspectual ambiguity

Turning to the mechanisms of ambiguity resolution in cases of ambiguous Aktionsart (Altmann & Steedman 1988), eyetracking during reading (e.g. Garnham, et al. (1997)), eyetracking in the visual world paradigm (e.g. Tanenhaus, et al. (1996)) and evoked potentials (e.g. van Berkum et al. (1999a)). A striking piece of evidence comes from van Berkum et al. (1999a) who found that the waveforms for one-referent- versus two-referent-contexts already lead to differences within 280 ms after the onset of the critical head noun. This shows that contextual information is used immediately during online interpretation. Thus referential contextual information appears to play an important role in the highly incremental process of determining initial preferences in syntactic disambiguation, too.

To sum up, probabilistic information like lexical frequency and contextual information show immediate effects in syntactic parsing. If the parser encounters an ambiguity concerning the subcategorization frame of a verb it immediately exploits these two kinds of information and chooses the subcategorization frame that fits the lexical and contextual requirements best.

4.3.1 The probabilistic parsing hypothesis

The first hypothesis, probabilistic parsing, stated in (7) assumes a parallelism between syntactic and aspectual semantic processing.

(7) Probabilistic Parsing Hypothesis (PPH)

Aspectual properties of a sentence are immediately determined on the basis of probabilistic information such as lexical aspectual frequency and context information.

I will use the example in (8) to illustrate how Aktionsart is processed under the PPH.

Hans built for two years a house
b. Hans baute in zwei Jahren ein Haus.
Hans built in two years a house

The PPH predicts that *baute* (built) will initially be interpreted as an accomplishment. This is due to a bias towards an accomplishment reading of *baute* in lexical frequency information. When the next segment is encountered, the *for-adverbial* in (8a) forces...
The lazy parsing hypothesis

According to the PPH both syntactic and aspectual semantic processing proceed in an incremental fashion and rely on probabilistic constraints. Here are the most important advantages and backdraws of the PPH:

- The PPH allows the incremental construction of temporal profiles. It is consistent with the results of Experiment 1 and the findings of other psycholinguistic experiments on aspectual coercion that found difficulty (like Piñango et al. (1999); Todorova et al. (2000a) and Husband et al. (2006)).
- Although being incremental, parsing along the lines of the PPH minimizes the amount of misrepresentation while processing local aspectual ambiguity. This is because in case of ambiguity always the most probable Aktionsart is chosen.
- The PPH is architecturally parsimonious. The syntactic and the aspectual semantic processor might even be the same device.

But the PPH also has some disadvantages:

- Due to immediate interpretation the PPH bears the risk of local misrepresentations. Thus even in a relatively innocent looking sentence like John built houses aspectual reanalysis is expected.
- Keeping track of aspectual semantic properties of sentences is a necessary prerequisite of the PPH (e.g. determining frequency distributions of the different Aktionsarten). This might be computationally costly, since differences in Aktionsart can often only be determined on the basis of very subtle semantic distinctions. Given the complexity of the task, it might be the case that the sentence processor is simply not able to determine the Aktionsart and fails to gain access to the lexical frequency information in aspectual lexical semantics. For example, the standard linguistic tests for Aktionsart that have been proposed by Vendler (1967) and Dowty (1979) heavily rely on semantic inferences like the following: does from x was φ-ing follow that x has φ-ed? Given the complexity of these tests, a quite natural assumption is that during aspectual semantic processing the parser is not able to fully determine the Aktionsart of the sentences.

4.3.2 The lazy parsing hypothesis

The lazy parsing hypothesis forms an intermediate step between probabilistic parsing and delayed aspectual processing. While aspectual semantic parsing is assumed to proceed incrementally the lazy parsing hypothesis presumes the aspectual semantic

processor not to have probabilistic information at its disposal. The lazy parsing hypothesis from Chapter (3) is repeated in (9).

(9) Lazy Parsing Hypothesis (LPH)

Every lexical item carrying aspectual information leads to a minimal update of the current Aktionsart such that it adds only as much to the model as is necessary to incorporate it into the semantic representation.

Again, it will be helpful to look at an example to illustrate how disambiguation works if the LPH is correct. For this purpose consider Example (8) repeated in (10) once again.

(10) a. Hans baute zwei Jahre lang ein Haus.
   Hans built for two years a house

b. Hans baute in zwei Jahren ein Haus.
   Hans built in two years a house

When the processor encounters baute (built) it immediately assigns an Aktionsart to the sentence fragment. Since build can be used as an activity Hans baute (Hans built) will receive an activity reading. On the next segment, in (10b) a type conflict emerges. This mismatch is resolved by transforming the activity into an accomplishment via additive coercion. Thus in (10b) we expect processing difficulty. When the processor eventually integrates the direct object ein Haus (a house), in (10a) it is taken as the incremental theme of a building activity; whereas in (10b) it serves as incremental theme (the continuously changing house) and simultaneously provides the culmination (the completion of the house), thereby specifying the missing parts of the accomplishment.

The verb bauen is initially interpreted as an activity, since this leads to the minimal update of the aspectual semantic representation, namely adding a preparatory process to it and nothing else. Note, however, that in most cases this update will lead to a later revision of the Aktionsart, since bauen (build) is most frequently used as an accomplishment. Thus, lazy parsing will very often lead to local misrepresentations with this particular kind of verb. If appropriate context information is available, the situation may be different. If it already provides a culmination the LPH predicts a different assignment of Aktionsart: in this case bauen can be given an accomplishment reading right from the start.

Again this hypothesis has a series of advantages, but also disadvantages which will be discussed in turn. First the advantages:

- Given its incremental nature, The LPH can explain the experimental findings of Experiment 1 and other psycholinguistic studies that reported difficulty (like e.g. Piñango et al. (1999); Todorova et al. (2000a) and Husband et al. (2006)).
For instance, according to the LPH the parser introduces a single event when it encounters the semelfactive verb sneeze in a sentence like John sneezed for an hour. When the verb is modified by the for-adverbial the semantic representation will be updated with a process by iterating the event into a repetitive process. This leads to immediate processing difficulty compared to a non-iterative control sentence.

- A second advantage of the LPH is that it can avoid some peculiarities that have been mentioned in connection with different kinds of direct objects like count versus mass nouns. This point is illustrated by the two sentences in (11).

(11) a. John drank a glass of water.
    b. John drank water.

While drink a glass of water in (11a) is an accomplishment, (11b) expresses an activity. The critical question is if one of these two forms is derivative of the other. Depending on the lexical frequency of the verb the PPH predicts that this is the case. Assume that drink is most frequently used as an accomplishment and John drank is interpreted accordingly. The bounded direct object in (11a) fits with the verb and should therefore be easy to process. By contrast, the mass noun object in (11b) triggers subtractive coercion. Intuitively, however, both sentences are easy to comprehend. The LPH gives an explanation for this: in both cases John drank introduces a process into the semantic representation. When the count NP is encountered in (11a) an incremental theme (= a glass of water), a culmination (= a finishing event) and the consequent state are introduced into the representation. There is no coercion involved in this semantic update. In (11b) the mass NP water is only consistent with an activity reading of which water is interpreted as the incremental theme. Again, the update works straightforward and no coercion is involved. Consequently, the LPH predicts both sentences in (11) to be roughly equally difficult.

- Third, the LPH provides a rather elegant explanation for the incremental buildup of complex Aktionsart. When an accomplishment has to be constructed, it can be put together piece by piece.
- Fourth, the LPH uses only simple lexical entries. No probabilistic information about the Aktionsart associated with a lexical item is needed.
- Finally, the LPH is computationally parsimonious since the semantic representation receives only a minimal update.

Besides its merits of which computational efficiency is certainly the most important the LPH has also a severe drawback:

- Although being computationally efficient, parsing along the lines of the LPH leads to more errors than probabilistic parsing. This is because selecting the most probable Aktionsart based on a number of different sources of information guarantees to always choose the most probable parse. By contrast, informationally limited heuristics like the LPH will in some cases predict an Aktionsart that is less likely than an alternative Aktionsart. Therefore, the LPC is not optimal in terms of minimizing the risk of local misrepresentations.

4.3.3 The aspectual underspecification hypothesis

The third hypothesis, aspectual underspecification, has been assumed in the literature by Pickering et al. (2006) who investigated aspectual coercion in a series of experiments using both self paced reading and eyetracking during ordinary reading. They studied the processing of constructions involving the computation of iterative readings and found no indication of difficulty for coerced sentences compared to non-coerced controls (for different findings see e.g. Husband et al. (2006)). On the basis of these null effects they claim that aspectual properties remain underspecified during sentence processing. What is most striking about their results is that within the same experiments they found difficulty for sentences with complement coercion (e.g. the author began the book) compared to non-coerced constructions (e.g. the author wrote the book). In the coerced sentence the book has to be reanalyzed as an event (like writing the book) while in the control construction it is interpreted as an artefact – which is already contained in the ordinary lexical meaning of book. Thus, although their experiments were actually sensitive to enhanced processing load due to semantic type-shift, there appeared to be no processing difficulty in the aspectual coercion cases.

The Aspectual Underspecification Hypothesis is restated from Chapter 3 in (12).

(12) Aspectual Underspecification Hypothesis (AUH)

Semantic aspectual interpretation remains locally underspecified during normal comprehension.

Let’s once again go through the Example (13) to see how processing under the AUH differs from the preceding two hypotheses.

    Hans built for two years a house
b. Hans baute in zwei Jahren ein Haus.
    Hans built in two years a house

4. This point is implicitly raised for instance in Hamm & van Lambalgen (2005) who treat (11a) as an instance of additive coercion, but at the same time use (11b) as an example of subtractive coercion.
According to the AUH, in both Examples (13a) and (b) the parser computes the Aktionsart of the sentence only when the last segment ein Haus (a house) is encountered. Thus, both adverbials in (13a) and (b) shouldn't differ in difficulty. It would make no difference if instead of the verb baute another transitive verb was used. According to the AUH processing of the adverbial phrase would always be like in (13a) and (b).

If the AUH is correct, there is a major difference between syntactic and aspectual semantic processing. While the syntactic processor immediately attaches new input according to a preferred structure, the computation of the aspectual representation is delayed. In syntax, this option is not available because of the quite limited working memory capacity in humans (see e.g. Cowan (2005)). New input has therefore to be immediately integrated into the structural representation. In semantics, however, the input is already structured resulting in much less memory load. In the most extreme version of the AUH the processor might wait until the end of the sentence and only then determine its Aktionsart. This minimizes the amount of aspectual reinterpretation. Let's briefly consider the advantages of the AUH:

- First, the parser will make fewer errors when interpretation is delayed.
- Second, the AUH provides an elegant explanation for sentence wrap-up effects commonly observed in psycholinguistic experimentation (described for instance in Haberlandt (1994)). Sentence wrap-up effects, i.e. a general slowdown in reading time at the end of a sentence, might be (at least partly) caused by an aspectual update of an underspecified representation.
- Third, one class of theoretical semantic accounts of aspectual coercion have extensively made use of underspecified representations (e.g. Dölling (2003a); Egg (2005)). If the AUH is correct these representations could be employed in language processing, too. Hence, performance in aspectual coercion can be explained by already existing semantic underspecification accounts.

Although on first sight the AUH looks quite attractive, it also has some obvious disadvantages.

- Delayed interpretation leads to higher working memory load than immediate interpretation, even though this effect might be less dramatic than in syntax. Memory load is enhanced because it is required that the lexical aspectual information of each word is stored until it can be eventually integrated. Thus, in terms of storage economy, the AUH is disadvantageous compared to immediate accounts of aspectual processing.
- The AUH is rather counterintuitive if one looks at instances of coercion. In these examples one has the intuitive feeling of difficulty while reading the sentence. Especially in sentences with multiple coercions one has the feeling of difficulty adding up. Under the AUH it would be predicted that difficulty may only emerge when full specification is carried out. This should probably only happen at the end of the sentence.
- In general, under the AUH, processing difficulty due to coercion is rather unexpected. Since at the end of a coerced sentence all the aspectual information is available, it should in principle cause no problem to determine its correct Aktionsart. If this is the case, the findings of Experiment 1 already provide evidence against the AUH. But two caveats are in order here: (1) the results of Experiment 1 were not obtained in ordinary reading, but might be due to the additional decision component in the incremental makes sense judgments and (2) processing difficulty was only found after the verb had already received all its arguments, so in some sense coercion effects were obtained only at the end of the experimental sentences which is fully consistent with the AUH.

4.3.4 A short summary

In the last sections I have been developing hypotheses about how ambiguity in aspectual semantics might be resolved during sentence processing. To this end, I first looked at disambiguation in syntactic processing which has been extensively studied in psycholinguistics. These studies suggest that the syntactic parser makes use of at least two sources of probabilistic information, lexical frequency and context information. However, with regard to aspectual ambiguity up to now basically nothing is known about the processes involved in disambiguation. I formulated three alternative hypotheses: the PPH which expects both syntactic and aspectual semantic disambiguation to rely on the same kind of probabilistic information, the LPH employing an incremental minimal update heuristic of the semantic representation and finally the AUH which assumes an aspectually underspecified semantic representation.

In the next two sections two experiments will be described which study how the processor deals with aspectual semantic ambiguity in online interpretation. The first investigates the role of lexical aspectual frequency while the second tests the influence of contextual information.

4.4 The role of lexical frequency information: Experiment 2

The first reading time study, Experiment 2, investigates the role of lexical frequency information in aspectual processing. Consider the initial parts of the German sentences in (14).
In these sentences two verbs are used that differ in terms of the Aktionsart which they predominantly express. While 
\textit{werkeln} (something in between \textit{potter about} and \textit{do handicrafts}) is almost exclusively used in sentences that express an activity, \textit{bauen} (\textit{build}) is most frequently used in accomplishment sentences. These preferences, however, are by no means absolute. An accomplishment use of \textit{werkeln} is illustrated in (15) which I googled from the internet.

(15) Hoffentlich lässt Electronic Arts die dann auch zu Ende werkeln.
Hopefully lets Electronic Arts them then also to the end potter about.

Hopefully, EA will let them finish their work on the video game.

Analogously, sentence (16), again from the internet, provides an example of an activity use of \textit{bauen} (\textit{build}).

(16) Sully baut derweil immer wieder an dem Haus herum.
Sully potters meanwhile always again at the house about
Right now, Sully potters about the house from time to time.

Potential difficulty on the adverbials in (14) allows us to decide between the theoretical options. The AUH predicts no difference in difficulty between for- and in-adverbials irrespective of the aspectual preference of the verb. Therefore, the adverbials in all conditions should not differ in reading time. The LPH predicts additive coercion in those sentences that contain an\textit{ in-adverbial} (b) and (d) while no coercion is involved in (a) and (c). Thus, longer reading times are expected in (b) and (d) compared to (a) and (c) irrespective of the lexical aspectual frequency of the verb. Finally, the PPH predicts additive coercion in (b) and subtractive coercion in (c), but no coercion in (a) and (d). Thus according to the PPH longer reading times are expected in (b) and (c) compared to (a) and (d).

4.5 Determining lexical frequencies: A corpus study

Determining the frequency distribution of the subcategorization frames of a verb is quite easy. One only needs a syntactically annotated corpus like the Penn Tree Bank and count how often a verb is used with a particular frame. For Aktionsart, this is not so simple. In Chapter 2 I described some of the difficulties in aspectual classification. First, to determine the Aktionsart with which a linguistic expression is associated one has to include the whole intrasentential and often also the extrasentential context. This makes the task extremely difficult for automatic or semiautomatic parsing. Second, although there exist a number of linguistic tests for Aktionsart, none of them is watertight due to the possibility of coercion. Thus, only converging evidence from a number of tests allows us to properly distinguish among the different Aktionsarten. Finally, many of these tests rest upon notions which are at the heart of sentential semantics like entailment, ambiguity with scopal operators and so forth (see the Vendler/Dowty criteria in Chapter 2). This implies that the linguistic tests crucially rely on intuitive semantic judgements and cannot be performed automatically.

As a result, to gather frequency information about aspectual preferences one has to ground the classification upon semantic judgments in a number of tests concerning whole sentences studied in their appropriate discourse context. This was done in a corpus study which I will now describe in more detail.

4.5.1 Method

The aim of the corpus study was to collect 20 verbs that are predominantly used to express activities as well as 20 accomplishment verbs. Two raters’ aspectually classified random samples from all written corpora of the Cosmas corpus (see their webpage at http://www.ids-mannheim.de/cosmas2/). The Cosmas corpus is a one-billion-word collection of connected discourse samples from more than four million written German

5. Thanks a lot to Andreas Konietzko with whom I spent never ending weeks over the corpus analysis.
texts including different genres (newspapers, novels, non-fiction, academic, technical reports etc). Most of the texts were published between 1980 and 2000.

First, about 40 activity and 40 accomplishment verbs were selected as candidates for the corpus study. All of them syntactically licensed an intransitive use (sentences of the form er verbt gerade (he is verb-ing right now) were well-formed). After a first inspection of corpus data for these verbs 20 activities and 20 accomplishments were selected. For each verb in its preterite form 100 hits were randomly selected from the corpus. Thereby, only hits were included that did neither contain idiomatic expressions (e.g. mir gehts gut (= I’m o.k.)) nor modals or constructions for which it was dubious what Aktionsart they have. Also, only cases for which both raters had the same opinion were included into the analysis. Because the verbs showed overall differences in frequency, two out of the 40 verbs ended up with less than 100 cases (70 and 95). Absolute sample sizes varied between 70 and 39621 hits. Each example was exported from the corpus together with the preceding context of five sentences.

The examples were classified in accordance with the linguistic tests from Dowty (1979) into the classical four aspectual classes: activities, accomplishments, achievements and states. This allows to compute a frequency distribution in percent for every verb. An example is provided for the activity verb backen (bake) and the accomplishment verb beladen (load) in Table 4.1.

Table 4.1 Frequency distribution for backen and beladen (absolute values)

<table>
<thead>
<tr>
<th>verb</th>
<th>activity</th>
<th>accomplishment</th>
<th>achievement</th>
<th>state</th>
</tr>
</thead>
<tbody>
<tr>
<td>backen</td>
<td>61</td>
<td>20</td>
<td>00</td>
<td>19</td>
</tr>
<tr>
<td>beladen</td>
<td>17</td>
<td>76</td>
<td>4</td>
<td>3</td>
</tr>
</tbody>
</table>

4.5.2 Results and discussion

The frequency distribution of each of the 40 verbs is provided in Appendix D. Because one has to take into account both activity and accomplishment frequencies the classification was carried out using the following criterion: computing the difference %activity – %accomplishment those verbs were classified as activities which have a positive value (are more often used in activity than in accomplishment contexts) while those with a negative value were classified as accomplishment. Figure 4.1 shows these differences in percent of use for all 40 verbs.

6. Additionally, it was made sure that the verbs were lexically unambiguous except for their Aktionsart.
The Processing of Events

4.6 Frequency information: Evidence from reading times

This experiment examined the effects of lexical frequency on aspectual ambiguity resolution. Subjects read aspectually ambiguous verbs which were followed by either an in or a for-adverbial which disambiguated the Aktionsart. Two different groups of ambiguous verbs were used. The activity verbs have relatively high corpus frequency of activity uses while the accomplishment verbs are predominantly used in accomplishment sentences. These preferences are, however, not absolute. Activity verbs can also appear in accomplishment sentences like (17a) and accomplishment verbs can be used to express an activity like (17b).

(17) a. Peter tratschte in einem Nachmittag sämtliche Geheimnisse aus.  
    In one afternoon, Peter disclosed all secrets.

b. Peter montierte zwei Tage lang an der Maschine herum.  
    For two days, Peter pottered about the machine.

c. Peter tratschte in einem Nachmittag sämtliche Geheimnisse aus.  
    In one afternoon, Peter disclosed all secrets.

d. Peter montierte zwei Tage lang an der Maschine herum.  
    For two days, Peter pottered about the machine.

Three factors were crossed in a factorial design: verb class (activity vs. accomplishment), adverbial (in vs. for-adverbial) and continuation (activity continuation vs. accomplishment continuation). Both activity and accomplishment verbs appeared equally often in sentences that were disambiguated towards an activity reading and in sentences that turned out to be accomplishment sentences. The segment containing the adverbial is the critical region to test for processing difficulty due to coercion.

    Hans jogged for two hours in the park.  
    For two hours, Hans jogged in the park.

    *Hans jogged in two hours in the park.  
    *In two hours, Hans jogged in the park.

    Hans jogged for two hours through the whole park.  
    For two hours, Hans was jogging through the whole park.

    Hans jogged in two hours through the whole park.  
    In two hours, Hans jogged through the whole park.

Two conditions are semantically deviant. The in-adverbial needs an argument which is telic, but in (18b) and (19b) it is combined with an atelic verb-argument structure. This semantic deviance is even stronger in the case of an activity verb in (18b) than in case of an accomplishment verb in (19b).

    Hans constructed for two hours about the model.  
    For two hours, Hans was constructing the model.

    ??Hans constructed in two hours about the model.  
    ??In two hours, Hans was constructing the model.

    Hans constructed for two hours the new model.  
    For two hours, Hans was constructing the new model.

    Hans constructed in two hours the new model.  
    In two hours, Hans constructed the new model.

Two conditions are semantically deviant. The in-adverbial needs an argument which is telic, but in (18b) and (19b) it is combined with an atelic verb-argument structure. This semantic deviance is even stronger in the case of an activity verb in (18b) than in case of an accomplishment verb in (19b).

20 item pairs like the one in (18) and in (19) were constructed using the verbs from the corpus study. The verb was always used in its past tense form to guarantee that the in-adverbial can only be interpreted in a durative manner. A list of stimuli is provided in Appendix E. Additionally 104 filler sentences were included in the experiment. 40 of them belonged to an unrelated experiment. The filler sentences were of various Aktionsart and 26 of them were semantically ill-formed (resulting in a overall ratio...
of 4:1 of well-formed to ill-formed sentences). The items and the distractors were distributed over four lists according to a Latin square design such that each item appeared under only one condition in each list.

For each sentence in the experiment a simple yes/no-comprehension question was formulated. The items were always followed by the question “did the sentence make sense?”. To keep subjects from preparing an answer while still reading the sentence 40 of the filler sentences had content questions.

Participants
Thirty six students from Tübingen University (all native German speakers; 25 female, mean age = 24.2, ranging from 19 to 35 years) participated in the experiment. The subjects were naive to the purpose of the study. Each subject was paid 5 Euros for participation. The participants were randomly assigned to lists (nine subjects per list).

Procedure
This experiment used self-paced reading of sentences using the moving window technique (see Haberlandt (1994)). The participants were told that they should read as fast as possible. Each sentence was followed by a comprehension question. The experiment started with written instructions. These were followed by a practice block with ten sentences in which explicit feedback was provided. The practice items contained no aspectual violations. After the practice followed the experiment in four blocks with a randomized order of sentences. Items and fillers were distributed across blocks such that each block contained an equal number of conditions and question types.

The experiment was conducted using an IBM compatible PC using E-Prime experimental software (www.pstnet.com). Responses were collected using a PST response box. The participants were tested individually in a quiet room at Tübingen University. The experiment took about half an hour.

Data analysis
Two dependent measures were analyzed: reading times on the adverbial and “yes, makes sense”-answers for the comprehension questions. These two measures allow to determine if coercion took place. For a coerced sentence longer reading times on the coercing segment are expected than in a control. Moreover, if coercion was successful subjects’ responses are expected to be “yes, the sentence made sense”.

Reading times were corrected for outliers by replacing reading times longer than 2500 ms by a value of 2500 ms and times below 100 ms by 100 ms. Trimming affected less than 2% of the data.

For statistical analysis repeated measures analyses of variance (ANOVA) were computed. Reading times on the adverbial, the preceding verb and the subsequent region were analyzed in a by-subjects ANOVA with the within factors verb class (activity vs. accomplishment) and adverbial (for vs. in). The by-items ANOVA was computed with the within factor adverbial (for vs. in) while verb class (activity vs. accomplishment) was treated as between factor. “Yes”-answers were analyzed in a by-subjects ANOVA with the within-factors verb class (activity vs. accomplishment), adverbial (for vs. in) and continuation (activity continuation vs. accomplishment continuation). Again, in the item analysis verb class was treated as between factor. Planned comparisons were computed using paired t-tests, both by participants (t_r) and by items (t_i). To reduce the danger of type I errors, planned comparisons always used a Bonferroni corrected α.

4.6.2 Results
Each subject scored better than 80% correct on the content questions. Thus, participants read attentively:

End of sentence judgments
Table 4.2 presents average “yes” answers for each of the eight conditions.

Table 4.2 Judgments in Experiment 2

<table>
<thead>
<tr>
<th>“yes, makes sense” answers</th>
</tr>
</thead>
<tbody>
<tr>
<td>activities</td>
</tr>
<tr>
<td>for-adverbial/activity continuation</td>
</tr>
<tr>
<td>in-adverbial/activity continuation</td>
</tr>
<tr>
<td>for-adverbial/accomplishment continuation</td>
</tr>
<tr>
<td>in-adverbial/accomplishment continuation</td>
</tr>
<tr>
<td>accomplishments</td>
</tr>
<tr>
<td>for-adverbial/activity continuation</td>
</tr>
<tr>
<td>in-adverbial/activity continuation</td>
</tr>
<tr>
<td>for-adverbial/accomplishment continuation</td>
</tr>
<tr>
<td>in-adverbial/accomplishment continuation</td>
</tr>
</tbody>
</table>

% “yes” answers (plus standard errors of the mean) in Experiment 2.

Sentences in which an in-adverbial modifies an activity sentence (very strongly biased by an activity verb and an activity continuation) were overwhelmingly judged nonsensical. Similarly, if an accomplishment verb is disambiguated towards an activity reading it is judged as semantically marked when it is modified by an in-adverbial. All the other conditions were rated above 80% as sensible. The only further difference consisted in slightly less acceptance in those conditions in which an activity verb is disambiguated towards an accomplishment reading than in the other conditions. The latter were judged as perfectly sensible. ANOVAs revealed a significant three-way interaction between verb class, adverbial and continuation (F1 (1, 31) = 35.87; MSE = .55; p < .01; F2 (1, 38) = 11.84; MSE = .36; p < .01) due to the difference between verb
classes in acceptability for *in-adverbials* modifying activity sentences. Also, the two-way interaction between *adverbial* and *continuation* was highly significant ($F_1 (1, 31) = 104.53; \text{MSE} = 4.33; p < .01$; $F_2 (1, 1) = 87.92; \text{MSE} = 2.64; p < .01$). This effect is due to less acceptance for *in-adverbials* that modify sentences with an activity continuation irrespective of the verb class. Planned comparisons of differences between semantic acceptance between conditions are presented in Table 4.3.

Table 4.3 Comparisons between “yes, makes sense” answers in Experiment 2

<table>
<thead>
<tr>
<th>activities</th>
<th>by subjects</th>
<th>by items</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>for/adverbial</em></td>
<td>$t_1 (31) = 15.01^*$</td>
<td>$t_2 (19) = 11.77^*$</td>
</tr>
<tr>
<td><em>for</em> vs. <em>in</em></td>
<td><em>for/adverbial</em></td>
<td>$t_1 (31) = 5.00^*$</td>
</tr>
<tr>
<td><em>for/continuation</em></td>
<td><em>for/adverbial</em></td>
<td>$t_1 (31) = .84$</td>
</tr>
<tr>
<td><em>for</em> vs. <em>in</em></td>
<td><em>accomplishment</em></td>
<td>$t_1 (31) = 8.33^*$</td>
</tr>
<tr>
<td><em>for/continuation</em></td>
<td><em>accomplishment</em></td>
<td>$t_1 (31) = .13$</td>
</tr>
<tr>
<td><em>for</em> vs. <em>in</em></td>
<td><em>accomplishment</em></td>
<td>$t_1 (31) = .22$</td>
</tr>
</tbody>
</table>

Paired t-test results for six planned comparisons under a Bonferroni corrected $\alpha$ of .0083. Asterisks indicate significant differences.

Reading times

Figure 4.3 shows adverbial effects (reading times for *in-adverbials* minus *for-adverbials*) for each sentence region. Increased processing difficulty for *in-adverbials* compared to *for-adverbials* is reflected by a positive deflection in reading times. I will report statistical analyses of reading times for the verb region, the adverbial region and the subsequent region.

The verb. There were no significant differences between the conditions in this region (all $F_s < 1.2$).

The adverbial. Reading times on the adverbial region are presented in Table 4.4. *In-adverbials* following an activity verb took longer to read than *for-adverbials*. After accomplishment verbs, however, there was only a slight difference. ANOVAs revealed a significant interaction between *verb class* and *adverbial* ($F (1, 31) = 11.69; \text{MSE} = 64158.10; p < .01$; $F_2 (1, 38) = 5.15; \text{MSE} = 41206.18; p < .05$). No other effects were significant. Planned comparisons revealed that following an activity verb *in-adverbials* are read more slowly than *for-adverbials* ($t_1 (31) = 3.3; p < .01$; $t_2 (19) = 5.22; p < .01$). By contrast, following an accomplishment verb there was no reliable difference in reading time between the two adverbials ($t_1, t_2 < 1$).

Table 4.4 Reading times on the adverbial in Experiment 2

<table>
<thead>
<tr>
<th>reading times</th>
</tr>
</thead>
<tbody>
<tr>
<td>activities</td>
</tr>
<tr>
<td><em>for-adverbial</em></td>
</tr>
<tr>
<td><em>in-adverbial</em></td>
</tr>
<tr>
<td>accomplishments</td>
</tr>
<tr>
<td><em>for-adverbial</em></td>
</tr>
<tr>
<td><em>in-adverbial</em></td>
</tr>
</tbody>
</table>

Reading times in milliseconds (plus mean standard errors).

The subsequent region. Because there are considerable lexical differences between the continuations for the two verb classes they were treated separately. Paired t-tests were computed to investigate differences between the *for-* and the *in-* *adverbials* within each class. The t-tests revealed no differences, neither for the activity verb class ($t_1, t_2 < 1$) nor for the accomplishments ($t_1, t_2 < 1$).

4.6.3 Discussion

This experiment investigated the influence of aspectual lexical frequency information on the disambiguation of Aktionsart during ordinary reading. If a telic *in-adverbial* followed an atelic activity verb reading times were longer than in the conditions with an atelic *for-adverbial*. This difference only showed up with the activity verbs, but not with the accomplishments. The latter were read equally fast both with *in-* and *for-adverbials*. The difficulty in processing *in-adverbials* following an activity verbs is a case of local semantic incompatibility since given the right kind of continuation (namely an accomplishment continuation) the sentences were judged as semantically well formed.
In this experiment processing difficulty due to telicity was found immediately at the disambiguating adverbial. This indicates that just like syntactic structure Aktionsart is processed incrementally during close-to-normal reading. Verbs were used which lead to a temporary ambiguity between an atelic and a telic Aktionsart. The following adverbial disambiguated between these two readings and caused processing difficulty in case the verb preferentially expressed an activity. The semantic acceptability judgements show that the aspectual mismatch can eventually be resolved: sentences were judged as sensible even if they contained an aspectual type conflict. This is particularly interesting since the processing of local aspectual ambiguity provides a purely semantic counterpart to the extensively studied garden path constructions. Aspectual mismatch has to be taken as a local phenomenon which can undergo aspectual coercion.

Turning to the three hypotheses that have been formulated in Section 4.1 the findings strongly support the PPH. First, in case of a preceding activity verb more difficulty was found with in-adverbials than with for-adverbials. This difficulty is due to an aspectual mismatch when an atelic activity verb is combined with a telic adverbial. Thus, the verb already must have been aspectually specified. Second, processing difficulty was only found within the activity verb class. This clearly indicates that the aspectual frequency information of a verb is considered during online comprehension and is used to predict the most likely analysis. Since the same factor is at work both in the syntactic and in the semantic domain it is possible that both employ the same cognitive architecture.

The findings are inconsistent with the LPH. Lazy parsing predicts more difficulty with in-adverbials than with for-adverbials irrespective of verb class. Contrary to this prediction only difficulty for the activity verbs, but not the accomplishments was observed. Hence, in aspectual parsing the aspectual representation does not always receive a minimal update, but is updated in a way to meet the affordances of what is most likely to come further downstream the sentence.

Similarly, the results provide evidence against the AUH. Underspecification predicts equal difficulty on the adverbial region irrespective if the adverbial is telic or not. This was clearly not the case in this experiment. Therefore, contra the proposal of Pickering et al. (2006) the aspectual properties of a yet incomplete sentence are immediately specified.

Pickering et al. (2006) found no evidence for difficulty in aspectual coercion measuring eye movements during reading although eyetracking is a particularly sensitive method that can detect rather subtle differences in processing load. How can the different findings between their study and the present experiment be explained? There is a major difference between the studies. While they solely focused on the processing of iterative readings, Experiment 2 investigated difficulty in additive and subtractive coercion. The findings can easily be reconciled if abstract type shifting does not cause difficulty, but other types of coercion do. This possibility will be investigated in Chapter 5.

Comparing the results of the present experiment with those of Experiment 1 there seems to be a discrepancy between findings. In Experiment 1 both additive and subtractive coercion caused difficulty. In the present experiment, however, only additive coercion showed an effect in reading times. That is, coercing an activity verb into an accomplishment (like it is semantically required by an in-adverbial) slows down reading pace, but stripping off the culmination from an accomplishment verb (like it is forced by a for-adverbial) does not. Thus, the results from Experiment 1 and 2 seem to be in conflict with each other. In principle, there are two possible explanations for this.

First, it might well be that the two types affected different components in the incremental makes sense decisions of Experiment 1. Additive coercion might be difficult because the semantic representation has to be updated with an additional eventuality. This affects both incremental makes sense judgements and reading times. By contrast, subtractive coercion might show a sole influence on the decision component in the incremental makes sense task. If this is correct, no difficulty during close-to-normal reading is expected. This possibility will be further investigated in Chapter 5.

Second, activity and accomplishment verbs are fundamentally different in verb argument structure. Activity verbs normally are intransitive while accomplishments need a direct object in addition which serves as incremental theme. If coercion is only possible in case the verb has all its arguments, difficulty is expected for the activity verbs, but not for the accomplishments. If this was the case the adverbial in the accomplishment cases simply appeared too early to be able to show any effects due to subtractive coercion. I investigate this possibility in Chapter 6.

So far, only intrasentential information has been used to provide disambiguation. For syntactic processing discourse context plays a key role in determining structural preferences. In the next experiment I will investigate whether this factor influences aspectual processing, too.

4.7 The role of context information: Experiment 3

A central and influential idea among semanticists is that meanings are computed obeying the compositionality principle which states that the meaning of an utterance is a function of the meaning of its parts and of the syntactic rules by which these parts are combined. Carrying this over to semantic processing has important consequences. Since syntactic rules operate at the sentence level, language interpretation is expected to proceed in a two-step fashion. First, the meaning of a sentence is computed. In a second step, the sentence meaning is integrated with information from prior discourse.
For instance, in their blueprint of the listener, Cutler & Clifton (1999) assume that, based on syntactic analysis and thematic processing, utterance interpretation takes place first and integration into a discourse model follows later. This view is in opposition to one-step models in which knowledge about the context and the world are brought to bear immediately, by the same system that combines the meanings of individual words into a larger whole. Recently, Hagoort & van Berkum (2007) have presented a one-step model with broad empirical coverage.

Up to now, discourse influences on aspectual processing have received practically no psycholinguistic attention (for an exception see Baggio, et al. (2008)). Since tense and aspect really play their role at the discourse level (Kamp & Reyle 1993) aspectual processing is a particularly interesting test case for two- versus one-step models. Experiment 3 tests the influence of contextual information on aspectual processing. The discourses in (20) illustrate how discourse context may influence the initial choice of Aktionsart.

(20) telic context:
Half a year ago, John started to jog about ten miles every day. When he began he was quite slow, but now he is really fast.

atelic context:
Half a year ago, John started jogging every day. When he began he had to stop after a short time, but now he can run for a long time.

for-target:
Today John jogged for one and a half hours.
in-target:
Today John jogged in one and a half hours.

The telic context introduces a path argument (ten miles). From the context it is clear that whenever John engages in a jogging activity it stops when the contextually introduced endpoint has been reached. If the target sentence can be directly connected to the discourse context it is expected that jog will be interpreted as an accomplishment. Thus, a target sentence containing an in-adverbial should be as easy to interpret as a target sentence containing a for-adverbial.

By contrast, the atelic context doesn’t introduce a path argument, but establishes jogging as an atelic activity. Jog in the target sentence can only be interpreted as an activity and therefore easily combines with a for-adverbial. However, modification by an in-adverbial is impossible since there is no culminating event with which the adverbial can be unified.

In Experiment 3 activity sentences were embedded in telic and atelic contexts to investigate whether contextual information can immediately overwrite local sentential information during aspectual processing. Processing difficulty was measured at in- versus for-adverbials in the target sentence. A one-step processing model predicts immediate and strong context influences. In contrast, two-step models predict that the target sentences will initially be processed alike irrespective of the preceding context.

4.7.1 Method
Experiment 3 was a self-paced reading study which presented whole discourses phrase by phrase using the moving window technique.

Materials
24 German discourses were constructed like in (21) each in four versions. Slashes indicate segmentation. A complete list of items is provided in appendix F.

(21) telic context:
Seit einem halben Jahr/schwimmt/Hans/jeden Morgen/zwei Kilometer/in einem Hallenbad./Anfangs/benötigte er/noch eine knappe Stunde dafür,/aber/er wird/von Tag zu Tag/schneller.
(For half a year now, Hans swims two kilometers in a swimming pool every morning. When he started he needed almost an hour for this distance, but he is becoming faster and faster.)

atelic context:
Seit einem halben Jahr/schwimmt/Hans/jeden Morgen/in/einem Hallenbad./Anfangs/konnte er/sich/noch kaum über Wasser halten,/aber/mit jedem Tag/geht es/besser.
(For half a year now, Hans swims in a swimming pool every morning. When he started he was almost drowning, but he is getting better every day.)

for-target:
Heute morgen/schwamm/er/dreißig Minuten/lang.
(in-target:
Heute morgen/schwamm/er/in nur dreißig/Minuten.

The activity verbs were chosen in a way that the described activity could both be done faster and longer as an effect of practicing. The contexts were always constructed similarly:

telic: the first sentence introduces an agent who regularly engages in a telic activity (+path argument). The second sentence expresses that the agent has gotten faster in performing the activity.

atelic: The first sentence introduces an agent who regularly engages in an atelic activity (−path argument). The second sentence expresses that the agent got considerably better in performing that activity.
The target sentence was always of the same form: a short time ago, the agent verb-ed for x time vs. in only x time. Note that without context the target sentences with an in-adverbial are semantically ill-formed (as the sensicality ratings for the activities in Experiment 5 show).

The adverbials were presented as two segments. This was done to have a spillover region following the critical first part of the adverbial phrase. In- and for-adverbials were selected with respect to length. The first segments (in nur dreißig vs. dreißig Minuten) were of comparable length: in-adverbials had a mean of 13.5 characters and for-adverbials had a mean of 13.9 characters (t(23) = −1.14; p = .27). The second part of the in-adverbials were 7.6 characters long and the for-adverbials were 5.0 characters long (t(23) = 18.08; p < .01). To take the latter difference into account, statistical analyses will be reported which were computed using reading times per character. Note, however, that the crucial comparison is within the same adverbial dependent on contextual influences.

Additionally, 64 filler discourses were constructed. Of these 34 were nonsensical. These were either globally (N = 24) incoherent (e.g. in a discourse expressing a causal connection the cause was happening after the effect) or locally (N = 10) incoherent (e.g. somebody kicking a ball under the goal). Each item received a question that queried whether the discourse was sensible (yes vs. no).

The experimental items and and the distractors were arranged in four lists according to a latin square design.

**Participants**

Thirty two students from Tübingen University (all native German speakers, 15 female, mean age = 22.5, ranging from 20 to 26 years) participated in the experiment. The subjects were naive to the purpose of the study. Each subject was paid 7.50 Euros for participation. The participants were randomly assigned to lists (eight subjects per list).

**Procedure**

Contexts and target sentences were presented on three lines (one sentence per line) in the center of the monitor such that target sentences didn’t stick out. The procedure was similar to the previous experiment. After reading the last segment of a discourse a question querying sensicality was presented. There was no time limit to provide a judgment.

An experimental session started with written instructions. They were followed by ten practice trials in which explicit feedback was given (all practice items were unrelated to the experimental items). The experiment consisted of four blocks containing 22 discourses each. It was made sure that each block contained the same number of items in each condition and that the ratio of sensible and nonsensical fillers was kept constant across blocks. Discourses were presented in a pseudo-randomized order such that at least two fillers intervened between two experimental items. An experimental session lasted for approximately 45 minutes.

### 4.7.2 Results

**Sensicality judgments**

While the sensible fillers received 84.2% “yes, makes sense” judgments, the non-sensical fillers were judged as sensible in only 20.9%. This indicates that readers paid attention to discourse coherence.

Figure 4.4 shows the sensicality ratings for the experimental conditions. Activity sentences with an in-adverbial following an atelic context were judged as sensible in 26.0%, but when these sentences were preceded by a telic context acceptance was 65.6%. For-adverbials were overwhelmingly judged as sensible irrespective of the preceding context (74.5% after a telic context and 79.7% after an atelic context).

![Figure 4.4. Sensicality judgments (+95% confidence intervals) in Experiment 3](image)

ANOVA on %-yes-answers revealed a significant interaction between context and adverbial (F₁(1, 31) = 38.34; MSE = 1.61; p < .01; F₂(1, 23) = 43.81; MSE = 1.16; p < .01). In-adverbials that followed a telic context received more positive ratings than when they followed an atelic context. For-adverbials showed the opposite pattern. Beside the interaction, both main effects of context (F₁(1, 31) = 22.70; MSE = .95; p < .01; F₂(1, 23) = 13.78; MSE = .68; p < .01) and of adverbial (F₁(1, 31) = 75.82; MSE = 3.16; p < .01; F₂(1, 23) = 46.01; MSE = 2.38; p < .01) were significant. The latter two effects...
resulted (a) from for-adverbials receiving more positive judgments than in-adverbials and (b) from the conditions with a telic context receiving on average more positive judgments than the conditions with atelic contexts.

Response times differed across conditions. The telic in-condition was judged in 1531 ms (SD 1900), but only in 1913 ms (2329) when they followed an atelic context. For-adverbials were judged in 1642 ms (SD 1597) when they followed a telic context and in 1355 ms (SD 1652) when they followed an atelic context. In ANOVAs, these differences lead to a interaction that was significant by items ($F(1, 31) = 4.08$; $MSE = 2921145.92$; $p = .06$; $F(1, 23) = 5.79$; $MSE = 2459136.75$; $p < .05$). No other effects were significant. The nonsensical distractors yielded judgment times of 1232 ms (SD 1292) compared to 1252 ms (SD 1292) for the sensible ones.

**Reading times**

Figure 4.5 shows the reading times of the target sentence. For illustration purposes, these have been aggregated over the verb and the subsequent pronoun region. The two adverbial regions have also been collapsed into one. However, statistical analyses will be reported both for aggregated reading times and reading times which look at the two adverbial regions separately from each other.

Up to the adverbial, reading pace did not differ across conditions (all $F$s < 1). When readers encountered the adverbial phrase they slowed down when it was aspectually inconsistent with the preceding context (telic/in: 102.5 ms/char.; atelic/in: 106.9 ms/char.) compared to the two aspectually matching conditions (atelic/for: 86.9 ms/char.; telic/in: 90.1 ms/char.). ANOVAs computed on reading times of the whole adverbial revealed a significant interaction between context and adverbial ($F(1, 31) = 12.55$; $MSE = 8435.63$; $p < .01$; $F(1, 23) = 11.48$; $MSE = 5839.86$; $p < .01$). But neither the main effect of context nor the main effect of adverbial reached significance (all $F$s < 1.5).

Table 4.5 shows the mean reading times of the two adverbial regions. The first region consisted of in only x (e.g. in nur dreißig) versus x time (e.g. dreißig Minuten), the second contained the last word of the adverbial phrase. When readers encountered the first part of the adverbial the matching conditions were read faster than the non-matching conditions (telic/in vs. atelic/in and atelic/for vs. telic/for). At the following region the same pattern showed up, leading to more than a 40 ms/char. difference between the matching and the non-matching conditions.

<p>| Table 4.5. Reading times on the two adverbial regions in Experiment 3 |
|----------------------|------------------|------------------|</p>
<table>
<thead>
<tr>
<th></th>
<th>first region</th>
<th>second region</th>
</tr>
</thead>
<tbody>
<tr>
<td>in-adverbial</td>
<td>53.14 (1.76)</td>
<td>157.43 (8.21)</td>
</tr>
<tr>
<td>for-adverbial</td>
<td>50.82 (1.96)</td>
<td>244.14 (12.40)</td>
</tr>
<tr>
<td>atelic context</td>
<td></td>
<td></td>
</tr>
<tr>
<td>in-adverbial</td>
<td>56.09 (1.63)</td>
<td>198.83 (8.83)</td>
</tr>
<tr>
<td>for-adverbial</td>
<td>45.26 (1.16)</td>
<td>197.61 (10.05)</td>
</tr>
</tbody>
</table>

Reading times per character (plus mean standard errors) in ms.

ANOVA comparisons reading times of the first part of the adverbial revealed a significant interaction between context and adverbial ($F(1, 31) = 16.98$; $MSE = 579.35$; $p < .01$; $F(1, 23) = 5.13$; $MSE = 416.44$; $p < .05$). Besides the interaction, the main effect of adverbial was significant ($F(1, 31) = 14.18$; $MSE = 1383.89$; $p < .01$; $F(1, 23) = 16.11$; $MSE = 1134.65$; $p < .01$). The main effect was due to in-adverbials being read more slowly than for-adverbials. To compare context effects within the two kinds of adverbials two planned comparisons were computed. Paired t-tests revealed that for-adverbials were read faster when they followed an atelic context than when they followed a telic context ($t(31) = 3.11$; $p < .01$; $t(23) = 2.16$; $p < .05$). But the numerical difference of in-adverbials contingent upon the preceding context was not reliable ($t(31) = 1.74$; $p = .09$; $t(23) = .92$; $p = .37$).

At the following segment, ANOVAs yielded a significant interaction between context and adverbial ($F(1, 31) = 15.24$; $MSE = 61848.44$; $p < .01$; $F(1, 23) = 12.55$; $p < .01$; $F(1, 23) = 3.11$; $p = .08$).

...
Discussion

For subtractive coercion see the discussion in the next chapter.

Alternatively, the activity verb may initially be taken as an activity. If the parser independence of contextual influences.

rests upon a strong context dependency assumption while the second assumes initial 

both in atelic and in telic contexts.

allowed. They also provide evidence that

ments indicate that given an appropriate context modification by an

in

activity, an

the blue, if the context includes a path-argument which provides an upper bound to

interpretation of Aktionsart. Activity sentences were tested which were modified by an

in

-adverbial can felicitously be combined with it. The sensicality judg-

ments indicate that given an appropriate context modification by an in-adverbial is 

allowed. They also provide evidence that for-adverbials can modify activity sentences 

both in atelic and in telic contexts.

The rating data are consistent with two alternative explanations of which the first 

rests upon a strong context dependency assumption while the second assumes initial 

independence of contextual influences.

1. Given a telic context, the parser may interpret an activity verb like jog as an implicit accomplishment right from the start. Thus, if it is followed by an in-adverbial, the adverbial can be composed with the accomplishment straightforwardly. In contrast, if the implicit accomplishment is followed by a for-adverbial, subtractive coercion is needed to make composition possible at all. Since subtractive coercion can be expected to enhance processing difficulty, a for-adverbial should be more difficult to comprehend after a telic context than after an atelic context.

2. Alternatively, the activity verb may initially be taken as an activity. If the parser encounters an in-adverbial the semantic requirements of the adverbial have to be fulfilled which consist both of preparatory process and a culmination. In a telic context, contextual information can provide the culmination while the prepara-

tion is already present in the sentence. As a result, an in-adverbial modifying an

activity can be felicitously interpreted in a telic context, but is semantically odd following an atelic context. If, by contrast, the parser is dealing with a for-adverbial no difficulty is expected since it should easily combine with an activity. Thus with for-adverbials no difficulty should emerge irrespective of the preceding context.

The reading time results are only compatible with the first alternative. Both, in-

adverbials following an atelic context and for-adverbials following a telic context lead to enhanced processing difficulty compared to the matching conditions. This provides evidence that sentential information is immediately combined with information from preceding discourse such that an activity verb like jog can directly be interpreted as an accomplishment. The data are thus fully compatible with one-step models like the one proposed by Hagoort & van Berkum (2007). But since self-paced reading data don't let us decide about the initial analysis of the linguistic input the present experiment can only serve as initial evidence for immediate context influences on aspectual parsing.

What is further interesting about the findings is that the aspectual information of the adverbial phrase is combined with the sentential and contextual information even before readers have reached the end of the phrase. This is indicated by the difference in reading time at the first segment of for-adverbials dependent on the preceding context. This shows the predictive abilities of the processor.

So, what happens when an activity verb is processed in a telic context? In some sense we are dealing with a kind of anaphora resolution here. An activity verb like jog has to be linked to a contextually salient accomplishment scenario (jog five miles). To do so, two things are necessary. First, the aspectual representation must be of the right format to allow for an update with additional aspectual information. Note that this is in principle possible even if the scenario is yet abstract, that is, when the representa-

tion contains information about a culminating event as an event variable without say-

ing what the culmination actually is. Second, variables have to be unified with concrete eventualities from preceding discourse. On the basis of the present experiment it is impossible to decide if these two aspects of a contextually triggered shift in Aktionsart can also be separated in processing. I will leave it as an open question if there is also facilitation for in-adverbials following activity verbs even if the contextual information doesn't include the concrete eventuality that is needed for coercion.

4.8 General discussion and conclusions

In this chapter I have investigated what factors guide the parser’s access to Aktionsart in the case of aspectual ambiguity. I reported two reading time experiments which studied the influence of lexical frequency information of verbs and the role of discourse context.
Taken together, the experiments provide evidence for probabilistic parsing (the PPH) commonly assumed in syntax. Readers had difficulty in interpreting an in-adverbial that followed a verb that is ambiguous between an activity and an accomplishment interpretation, but only if the activity interpretation is more frequent with respect to the particular verb. Context information can override lexical preferences. If an activity verb is embedded in a telic context it is immediately interpreted as an accomplishment.

The results are inconsistent with the two alternative hypotheses. First, during ordinary reading there was immediate difficulty caused by aspectually mismatching adverbials. This finding speaks against aspectual underspecification (the AUH) inspired by Pickering et al. (2006). Second, the processor distinguished between different kinds of verbs that were ambiguous between activities and accomplishments taking lexical frequency information of Aktionsart into account. This is incompatible with lazy parsing (the LPH) which assumes a minimal update of the aspectual representation irrespective of probabilistic information.

To sum up, the present chapter provides evidence for immediate assignment of Aktionsart during language comprehension. But the initial aspectual analysis may be in conflict with aspectual information that comes later in a sentence or a discourse. The question arises how and when reanalysis is performed. This will be the topic of the following chapters.

CHAPTER 5

Processing different types of coercion

In this chapter three different types of aspectual coercion will be studied. I will report three self-paced reading experiments which investigated processing of additive coercion (Experiment 4a), subtractive coercion (Experiment 5) and abstract type shift (Experiment 6). The experiments provide evidence that only additive coercion causes difficulty. This suggests that it is abductive reasoning about missing aspectual information that is cognitively demanding.

The different kinds of coercion allow us to investigate the processing consequences of fundamental semantic properties of natural language: context dependency, non-monotonicity and flexible type-shifting.

Strict compositionality states that the meaning of an utterance is a function of the meaning of its parts and of the syntactic rules by which they are combined (e.g. Partee (1984)). In additive coercion, however, the sentence meaning has to be integrated with information from prior discourse or world knowledge. This is illustrated by the two discourses in (1).

(1) a. A mountaineer was hiking to the basecamp. He reached the camp in five hours.

b. A mountaineer got on the helicopter which was leaving the airport for the basecamp. He reached the camp in five hours.

The interpretation of the second sentence depends on the discourse context provided by the first sentence. While the second sentence in (1a) means that the mountaineer reached the camp after a five hour long hike, the exact same sentence in (1b) means that he reached the camp after a five hour long helicopter flight. Thus, in (1) interpretation does depend on sentence external information and cannot be determined on the sole basis of the lexical items that are contained in the coerced sentence and the way they are combined. The lexical information of the achievement reach the camp is simply not sufficient to determine the meaning of reach the camp in five hours on the basis of the sentential information alone. Still, the aspectually enriched interpretation comes about by just adding information to the existing semantic representation without the need to revise it. Here, another fundamental semantic concept comes into play, namely monotonicity.
Non-monotonic logic covers formal frameworks devised to capture and represent defeasible inference, i.e. that kind of inference of everyday life in which reasoners draw conclusions tentatively, reserving the right to retract them in the light of further information. Such inferences are called “non-monotonic” because the set of conclusions warranted on the basis of a given knowledge base does not necessarily increase (in fact, it can shrink) with the size of the knowledge base itself. This is in contrast to classical (first-order) logic, whose inferences, being deductively valid, can never be “undone” by new information. Monotonicity states that if \( \varphi \) is a consequence of \( \Gamma \) then it is also a consequence of any set containing \( \Gamma \) as a subset.

Monotonicity: If \( \Gamma \vdash \varphi \) and \( \Gamma \subseteq \Delta \) then \( \Delta \vdash \varphi \)

In logic programming (which serves as the underlying framework of the event calculus of Hamm & van Lambalgen (2005)) non-monotonicity comes about by the closed world assumption which states that information that is not given or inferred is (temporarily) assumed to be false. The closed world assumption leads to the construction of minimal models, that is, the simplest possible structure which makes a narrative true. The computation and recomputation of a minimal model is exemplified by the unfolding piece of discourse in (2).

(2) a. Der Architekt baute ein Haus ...
   The architect built a house ...

b. Der Architekt baute ein Haus, bis die Bank den Kredit kündigte
   The architect built a house until the bank the credit canceled

Assuming incremental discourse interpretation, when (2a) is encountered, the processor will compute a minimal model in which the consequent state is attained, that is, in which a complete house is part of the discourse representation. This computation is based on the “closed world assumption”: because no disabling condition is described in the discourse, it will be (temporarily) assumed that there is no obstacle to attaining the goal. When the initial model is extended with a subordinate clause like (2b) describing an event which terminates the building activity (a disabling condition), the goal state inference will be suppressed. The clause when the bank canceled the credit will lead to the retrieval of causal knowledge from semantic memory to the effect that the credit failure accident probably terminated the building activity. Credit failure occurred during the building, so it follows that the accident took place before a complete house was obtained. Thus, (2b) requires recomputation. Non-monotonic update processes of discourse models are not restricted to the aspectual domain but are a rather common phenomenon. This is illustrated in (3).

(3) a. Some politicians are corrupt. In fact, all politicians are corrupt.
   b. In New York, a man gets mugged every hour. He was on TV last night.
   c. The triangle is to the left of the circle. The square is in between the two.

In (3a) involves canceling the implicature some \( x \) are \( y \) \( \rightarrow \) implicates not all \( x \) are \( y \) (Grice 1967), in (3b) the resolution of the anaphor he forces the recomputation of quantifier scope and in (3c) the spatial configuration of the objects has to be reconsidered. Very little is known about the processing consequences of non-monotonic update processes.

Flexible type shifting is another interesting property of natural language semantics. Intuitively, expressions like 
sneeze and the accelerator have one lexical meaning. The linguistic context in which they appear may however shift their interpretation. Consider (4).

(4) a. John sneezed for almost an hour.
   b. The accelerator is on the right of the brake.

In (4a) sneeze has to be type shifted from an event into a process. Most naturally this can be achieved via iteration yielding an interpretation where John sneezed over and over again. Type-shifting is not restricted to eventualities but can also occur within the nominal domain as illustrated in (4b). In order to get a generic interpretation of the sentence, the accelerator has to receive a type reading in contrast with the intuitively preferred token interpretation. Language interpretation may make excessive use of type shifting since it is commonly assumed that even innocent looking quantified sentences like The teacher praised every kid and coordinated structures like John and every man involve type shifting (e.g. Partee & Rooth (1983)). In the latter example John has to be shifted from a name of type e into a generalized quantifier (of type (e, t, t)).

To sum up, contextual enrichment, non-monotonic update and flexible type shifting are particularly interesting properties of semantic interpretation. Do they cause difficulty in processing? This will be the topic of the next three sections.

5.1 Additive coercion: Experiment 4a

The two experiments reported in the previous chapter already made use of constructions involving additive coercion. In both experiments an activity was coerced into an accomplishment by implicitly adding a culmination. We have thus already obtained evidence for additive coercion causing difficulty. However, the constructions used in these experiments were aspectually ambiguous. It is thus possible that difficulty was due to activating a dispreferred reading instead of constructing a new situation type. To obtain clear results about reinterpretation we need to investigate definite transitions from one aspectual class to another.

Unambiguous achievements that are coerced into accomplishments certainly involve aspectual reinterpretation. In this construction type an achievement is updated with an appropriate preparatory process which is finished by the culmination.
Experiment 1 already tested this construction type and provided offline evidence for difficulty at the coercing in-adverbial in sentences like (5b).

(5) a. Der Kletterer erreichte den Gipfel vor zwei Stunden.
The climber reached the top two hours ago.

b. Der Kletterer erreichte den Gipfel in zwei Stunden.
The climber reached the top in two hours.

The climber reached the top for two hours.

However, Experiment 1 did not compare additive coercion in (5b) to an aspectual control like (5a) which also involves an achievement. Using the three constructions in (5) allows to compare additive coercion (5b) to aspectual control (5a) and aspectual mismatch (5c). The latter comparison is interesting because it may tell us whether coercion adds something to mismatch detection. If coercion and mismatch show effects with a different time course this can tell us about functional differences between the two conditions. In particular, the reading time patterns might indicate that mismatch detection and repair are two processes that can be differentiated. While mismatch is expected to show a rather focused effect, repair is expected to lead to a later effect.

5.1.1 The aspectual enrichment hypothesis (AEH)

The AEH which is repeated in (6) predicts processing difficulty in additive coercion.

(6) Aspe c tual Enrichment Hypothesis (AEH)
Enriching the aspectual representation with some additional eventuality causes processing difficulty.

The AEH is a corollary of the semantic enrichment hypothesis (7) which attributes processing difficulty to semantic enrichment in general.

(7) Semantic Enrichment Hypothesis (SEH)
Enriching the semantic representation with additional content which is not explicitly stated in the sentence causes difficulty.

The SEH uniformly covers aspectual coercion and complement coercion. The latter is once more illustrated in (8a) with its most plausible interpretation (8b).

(8) a. The author began the book.

b. The author began writing the book.

It is obvious that in order to compute (8b) the reader must add semantic content which is not linguistically encoded in (8a). Analogously to additive aspectual coercion, world knowledge is needed to infer the right kind of activity which allows to type shift the book from an artefact into an event. If the sentence were for instance about a reader beginning a book, the appropriate activity would have been completely different.

Complement coercion has been extensively studied by Brian McElree’s group applying a wide range of methods (for an overview see Chapter 2). These studies provide solid evidence for processing difficulty in this type of coercion. If additive coercion also leads to processing difficulty this would add evidence for the general SHE also in the verbal domain.

5.1.2 Method

Experiment 4a was part of an experiment which tested two word order variants. In Experiment 4a the adverbial followed a complete verb-argument structure (SVOA-order). The results of the subject-verb-adverbial-object order (SVAO-order) will be reported separately in the next chapter (see Experiment 4b).

Materials

30 items like the one in (9) were constructed. All items used transitive achievement verbs (like win, reach, etc.) in the simple past. A list of stimuli is provided in Appendix G.

(9) a. Der Athlet/ gewann/ die Medaille/ vor drei Stunden/ bei den Para-Olympics.
The athlete won the medal three hours ago at the para olympics.

b. Der Athlet/ gewann/ die Medaille/ in drei Stunden/ bei den Para-Olympics.
The athlete won the medal in three hours at the para olympics.

The athlete won the medal for three hours at the para olympics.

Two factors were crossed in a factorial design: word order (SVOA vs. SVAO) and adverbial (ago vs. in vs. for-adverbial). The segment containing the adverbial was the critical region testing for processing difficulty due to aspectual coercion and aspectual mismatch.

Additionally 75 filler sentences were included in the experiment. They encompassed all kinds of Aktionsarten (11 states, 23 activities, 21 accomplishments, 9 achievements and 11 semelfactives) and 25 of them were semantically ill-formed resulting in an overall ratio of 2:1 of well-formed to ill-formed sentences. The items of experiments 4a and
4b and the distractors were distributed over six lists according to a latin square design such that each item appeared under only one condition in each list.

For each sentence a simple yes/no-question was formulated. The items were always followed by “did the sentence make sense?”. To keep subjects from preparing an answer while still reading the sentence 35 of the filler sentences had ordinary comprehension questions which asked for content.

Participants

Thirty students from Tübingen University (all native German speakers, 24 female, mean age = 22.9, ranging between 19 and 45 years) participated in the experiment. The subjects were naive to the purpose of the study. Each subject was paid 5 Euros for participation. The subjects were randomly assigned to lists (five subjects per list). Three additional subjects were excluded due to poor performance on the comprehension questions for the fillers.

Procedure

Experiment 4a used self-paced reading with moving window presentation just like Experiment 2 with the only exception that there was a time limit of five seconds to provide an answer (measured from the question onset).

The experiment started with written instructions. These were followed by a practice block of ten sentences in which explicit feedback was provided. The practice items contained no aspectual violations. After the practice followed the experiment in one block with a randomized order of sentences. The experiment took about twenty minutes.

Data analysis

Two dependent measures were analyzed: reading times on the adverbial and “yes, makes sense”-answers for the comprehension questions.

The reading times were trimmed by replacing reading times longer than 2500ms by a value of 2500ms and times below 100ms by 100ms. This correction affected less than 1.5% of the data. Because the adverbial phrase minimally differed in length across conditions reading times per character were used for the statistical analysis. An analysis of the rather short preposition revealed no differences between conditions suggesting that readers had skipped this segment. The last two segments were thus treated as one region in the analysis.

Repeated measures analyses of variance (ANOVAs) were computed. Reading times on the adverbial, the subsequent preposition and the rest of the prepositional phrase were analyzed in subjects and items ANOVAs with the within factor adverbial (control vs. coercion vs. mismatch). Similarly, “yes, makes sense”-answers were analyzed in subjects and items ANOVAs with the within-factor adverbial (control vs. coercion vs. mismatch). Planned comparisons were computed using paired t-tests, both by participants ($t_1$) and by items ($t_2$). To reduce the chance of type 1 errors, planned comparisons always used a Bonferroni corrected $\alpha$.

5.4.3 Results

“Makes sense” judgements

Participants judged the control condition as sensible in 93.2%, additive coercion in 48.5% and mismatch in 16.0%. In ANOVAs, this difference lead to a significant main effect of adverbial ($F_1 (2, 58) = 182.56; \text{MSE} = 5.41; p < .01; F_2 (2, 58) = 75.24; \text{MSE} = 5.06; p < .01$). Participants were quite homogenous in their judgment pattern in the coercion condition which is illustrated by the distribution of judgments in Figure 5.1.

Numerically, judgment times were longest after sentences involving additive coercion (1580ms) followed by control (1483ms) and aspectual mismatch (1410ms). However, this difference was not statistically reliable ($F_1 (2, 58) = 2.49; \text{MSE} = 230491; p = .10; F_2 (2, 58) = 1.74; \text{MSE} = 231477; p = .19$).

Reading times

Diagram 5.2 shows mean reading times of sentences involving coercion and mismatch compared to control for the whole sentence.

Up to the adverbial phrase the three aspectual conditions were identical and did not differ in reading time.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure5.1.png}
\caption{Distribution of judgments over participants}
\end{figure}
When readers encountered the adverbial phrase they slowed down in case of a *for-adverbial* (mean RT 60.0 ms/char.) and in case of an *in-adverbial* (mean RT 60.9 ms/char.) compared to aspectual control (mean RT 54.1 ms/char.). In ANOVAs, this difference was reflected in a significant main effect of adverbial ($F_1 (2, 58) = 4.05; \text{MSE} = 652.23; p < .05$; $F_2 (2, 58) = 3.49; \text{MSE} = 640.45; p < .05$). One-tailed planned comparisons revealed that mismatch was read more slowly than control ($t_1 (29) = 2.26; p < .025$; $t_2 (29) = 2.32; p < .025$) and that additive coercion was read more slowly than control ($t_1 (29) = 2.34; p < .025$; $t_2 (29) = 2.44; p < .025$).

Further, reading times of coerced sentences were analyzed contingent on judgments. Thus, those trials in which participants judged a sentence with an in-adverbial region (mean RT/char. 105.5ms), while mismatch (mean RT/char. 93.4ms) and control (mean RT/char. 88.9ms) were roughly the same. In ANOVAs, this difference resulted in a significant main effect of adverbial ($F_1 (2, 58) = 3.04; \text{MSE} = 2281.09; p < .05$; $F_2 (2, 58) = 3.49; \text{MSE} = 640.45; p < .05$). One-tailed planned comparisons revealed that mismatch was read more slowly than control ($t_1 (29) = 2.26; p < .025$; $t_2 (29) = 2.32; p < .025$) and that additive coercion was read more slowly than control ($t_1 (29) = 2.34; p < .025$; $t_2 (29) = 2.44; p < .025$).

Figure 5.2. Reading times phrase by phrase in Experiment 4a. Error bars represent 95% confidence intervals of the mean.

In the additive coercion condition the slow-down extended to the subsequent PP region (mean RT/char. 105.5ms), while mismatch (mean RT/char. 93.4ms) and control (mean RT/char. 88.9ms) were roughly the same. In ANOVAs, this difference resulted in a significant main effect of adverbial ($F_1 (2, 58) = 7.56; \text{MSE} = 2711.68; p < .01$; $F_2 (2, 58) = 4.83; \text{MSE} = 2281.09; p < .05$). One-tailed planned comparisons revealed that reading times in the coercion condition were slower than in the mismatch condition ($t_1 (29) = 3.04; p < .025$; $t_2 (29) = 2.42; p < .025$). There was, however, no significant difference between mismatch and control ($t_1 (29) = 1.21; p = .24$; $t_2 (29) = .78; p = .44$).

Further, reading times of coerced sentences were analyzed contingent on judgments. Thus, those trials in which participants judged a sentence with an in-adverbial...
in successfully coerced sentences as if they were checking the following PP for the required information.

Comparing the results of the present experiment with psycholinguistic studies on complement coercion (e.g. McElree et al. (2001); Traxler et al. (2002)), the data lend support to the view that semantic enrichment enhances processing demands in general (the SEH). Complement coercion involves type shifting of an argument into an event (for example of book into writing a book) by inferencing the right kind of process on the basis of pragmatic information; it provides an instance of additive coercion in the nominal domain. The present experiment shows difficulty in additive coercion in the verbal domain. Taken together, there is converging evidence that the pragmatic strengthening of an impoverished semantic representation causes difficulty independently of the concrete phenomenon under investigation.

As expected, aspectual mismatch was rejected across the board and lead to longer reading times at the mismatching adverbial. But there is a striking difference between the reading times in the mismatch and the coercion conditions. In the mismatch condition, subjects detected the aspectual violation right away and read the following region with normal reading pace comparable to control. In coercion, the effect was more sustained. When comprehenders read coercing adverbials, but were not able to come up with the right kind of eventualty reading pace was reduced including the last region. This indicates that in additive coercion readers did not reject the sentence as nonsensical when they encountered the coercing stimulus but tried to repair the aspectual representation way down the sentence.

The present experiment investigated a type of coercion which hasn't up to now been studied up to now in the psycholinguistic literature. The findings show that additive coercion causes difficulty when no appropriate context is available. In a next step, it will be interesting to look into the functional processes in additive coercion. I will come back to this issue in Chapters 6 and 7.

5.2 Subtractive coercion: Experiment 5

In this section I will report a reading time study which investigated subtractive coercion. A complex Aktionsart is coerced into a more simple Aktionsart by stripping off some part of the eventualty. It involves non-monotonic update processes, that is an inference that was valid before does not have to hold anymore after the Aktionsart has been coerced. In line with the Non-Monotonic Update Hypothesis, we may thus expect processing difficulty in subtractive coercion.

(10)  Non-Monotonic Update Hypothesis (NMUH)

The revision and recomputation of a discourse model causes processing difficulty.

Monotonocity has been discussed in the psycholinguistic literature in relation to "reanalysis" in syntactic processing (Fodor & Ferreira 1998). If a locally ambiguous sentence is disambiguated towards the dispreferred reading, the preferred structure has to be revised and reanalyzed. This implies that the syntactic attachments which were established in the first parse must be discarded, which contradicts monotonocity. It is well established in syntactic processing that non-monotonic updates of sentence structure cause difficulty, namely the well known garden-path effects.

Recomputing a discourse model involves semantic reanalysis of the initial discourse representation. Semantic information has to be canceled just like wrong syntactic attachment has to be revised in repairing a syntactic garden-path. It is thus quite plausible to assume the NMUH and to expect difficulty in processing sentences involving subtractive coercion.

5.2.1 Method

Experiment 5 used self-paced reading to investigate subtractive coercion in sentences like (11).

(11)  a. Der Bergsteiger bestieg den K2 in zwei Tagen trotz eines Sturms.
    The mountaineer climbed the K2 in two days in spite of a storm

b. Der Bergsteiger bestieg den K2 zwei Tage lang trotz eines Sturms.
    The mountaineer climbed the K2 for two days in spite of a storm

The initial part of the sentence Der Bergsteiger bestieg den K2 is an unambiguous accomplishment and entails that the mountaineer reached the top of K2. When continued with ohne oben anzukommen (without reaching the top) the sentence becomes contradictory. Interestingly, the contradiction disappears if a for-adverbial intervenes. Consider (12).

(12)  a. Der Bergsteiger bestieg den K2 zwei Tage lang.
    The mountaineer climbed K2 for two days
    ohne oben anzukommen.
    without reaching the top.

Here, the for-adverbial shifts the accomplishment into an activity. This is a clear instance of subtractive coercion involving a non-monotonic update of the aspectual representation.

Materials

20 items were constructed using unambiguous transitive accomplishment verbs. The verbs were carefully chosen because for instance ein Haus bauen (build a house), which is the prototypical example of an English accomplishment can easily receive an activity reading in German. To put the potential accomplishments to test, all verbs
plus their direct objects were put into the sentence frame in (13). Only VPs which led to a contradiction (= consistently judged by two informants) were considered to be accomplishments. Further, only verbs were selected which were not homonymous according to the Duden lexicon of German (leaving out verbs like *erbauen* which can either mean build or elevate).

(13) John [VP_{simple-pas}] but he didn’t finish.

The items were constructed in four conditions according to a factorial design including the factors adverbial (for vs. in) and word order (verb-object-adverbial (VOA) vs. verb-adverbial-object (VAO)). The word order variation was included to investigate if difficulty in subtractive coercion can emerge before the direct object has been read, that is, before the incremental theme is specified. (14) is a sample item (see Appendix H for a full list of items). Slashes indicate segmentation.

   Hans built the house in two years in spite of financial problems.

   Hans built the house for two years in spite of financial problems.

c. Hans/ errichtete/ in zwei Jahren/ trotz/ finanzieller Probleme/ das Haus...
   Hans built in two years in spite of financial problems the house...

d. Hans/ errichtete/ zwei Jahren lang/ trotz/ finanzieller Probleme/ das Haus...
   Hans built for two years in spite of financial problems the house...

In the VAO word order in (14c) and (14d) an additional PP was added at the end of the sentence to prevent interference of potential effects at the postponed direct object with sentence wrap-up.

Each item received a question which either queried if the sentence expressed that the preparatory process had been completed (*wurde fertig ge-verb-t?*) in case of the first 12 items or if the sentence made sense for the last eight items.

48 fillers were included from Experiment 1 which were slightly modified (40 sensible, 8 not sensible). Further, a second experiment testing 16 activity verbs under for- and in-adverbials as in (15) was also included in the experiment. With the activities, an in-adverbial is expected to introduce difficulty, which is the exact opposite pattern than with the accomplishments. All filler trials received a question, which either queried if the sentence was perceived as sensible (N = 32) or asked for content (N = 32). Four lists were constructed according to a latin square design.

   the runner ran for ten minutes in the new stadium.

b. Der Sprinter/ lief/ in zehn Minuten lang/ im neuen Stadion.
   the runner ran for ten minutes in the new stadium.

Additionally 10 practice items were formulated which did not contain any asp-ec-tual violation.

Participants
32 native German speakers recruited at Tübingen University participated in the experiment (24 female, mean age = 22.8, ranging from 19 to 31). All participants were naive to the purpose of the study. Each subject was paid 5 Euros for participation. The participants were randomly assigned to lists (eight subjects per list). Two additional subjects were excluded due to poor performance on the comprehension questions for the fillers.

Procedure and statistical analysis
The procedure was identical to Experiment 4a. The experiment took approximately 25 minutes.

Like in Experiment 4a, for purposes of statistical analysis reading times per character were analyzed. Answers to the two types of questions (attainment of the goal state vs. sensibility ratings) were analyzed separately.

5.2.2 Results
Judgments
Figure 5.3 shows the judgment results for both kinds of questions.

Accomplishments modified by a for-adverbial were more often judged as incomplete than accomplishments modified by an in-adverbial. This difference was more pronounced in the canonical VOA order (for: 56.6% vs. in: 86.4%) than in the VAO order (for: 67.2% vs. in: 80.1%). Accomplishments were overwhelmingly judged as sensible irrespective of the kind of adverbial (for: 69.5% vs. in: 71.0%). Participants clearly noticed asp-ectual violations. This is indicated by the sensibility judgments provided for activities (for: 88% vs. in: 31%).

Statistical analyses on attainment of the goal state-judgments revealed a signifi-cant main effect of adverbial ($F_1(1, 31) = 14.13, MSE = 1.35, p < .01$; $F_2(1, 11) = 34.04, MSE = 0.58, p < .01$). This main effect is due to participants choosing a “no, not completed” answer more often when they were dealing with a for-adverbial than when they were dealing with an in-adverbial. Neither the main effect of word order nor
the interaction between *adverbial* and *word order* reached significance. Planned comparisons were computed which contrasted *for-* with *in-*judgments separately for the two word orders. They yielded a reliable difference in case of VOA order ($t_1(31) = 4.07, p < .025$; $t_2(11) = 3.89, p < .025$) but only a marginally significant difference in case of VAO order ($t_1(31) = 2.04, p = .05; t_2(11) = 2.42, p = .034$). Thus, the type-shifting effect of *for* adverbials was bigger in the unmarked VOA order than in the marked VAO order. The latter contained two additional segments making it however possible that effects were covered up at the end of the sentence.

Two planned comparisons of *makes sense* judgements compared accomplishment and activity sentences with *for-*adverbials to those with *in-*adverbials. In terms of semantic acceptability, there was no difference between subtractive coercion and control ($t_1(31) = 0.32, p = .75; t_2(7) = 0.09, p = .93$). However, in the activity sentences there was a big difference between the two types of adverbials ($t_1(31) = 11.20, p < .01; t_2(15) = 12.97, p < .01$).

ANOVAAs were computed which analyzed judgment times for the accomplishments. There were no significant differences between conditions.

**Reading times**

The left-hand side of Figure 5.4 depicts the mean reading times for subtractive coercion and control sentences in the canonical VOA order. Throughout the sentence reading times did not differ between subtractive coercion and control (all Fs $< 1$). At the critical adverbial region, subtractive coercion was numerically even faster than control. In the VAO order subtractive coercion and control also didn’t show a difference (all Fs $< 1$). Reading times in the VAO order are depicted on the right-hand side of Figure 5.4.

Activities under *in* and *for* modification, however, did differ in reading times at the adverbial. They are depicted in Figure 5.5.

Paired $t$-tests yielded a significant difference in reading time of the adverbial phrase ($t_1(31) = 4.30, p < .01; t_2(15) = 6.98, p < .01$) resulting from the aspectually inappropriate *in-*adverbials being slower than the *for-*adverbials.

### 5.2.3 Discussion

The present experiment investigated whether perfective sentences involving subtractive coercion cause processing difficulty. Although the judgment data suggest that readers coerced accomplishments into activities when the sentence contained a *for-*adverbial, reading times did not differ across conditions. In the VOA order reading times at the coercing adverbial were numerically even faster than in the control condition. Thus, subtractive coercion does not cause any measurable disruption in reading.

Before rashly discarding the NMUH there is an incongruous finding which calls for an explanation. In the incremental makes sense judgements in Experiment 1 subtractive coercion lead to longer decision times in the exact same constructions that were used in the present experiment. The only difference between the experiments consists in the employed method. How can the opposite findings be reconciled? One way to proceed is to look at alternative ways a subtractively coerced meaning can be computed and to investigate how cognitively demanding they are. There are two different computations which both compute the same meaning but differ algorithmically. The actual computation is implemented using simple recurrent neural nets. The underlying framework has been developed by Stenning & van Lambalgen (2005) and van Lambalgen & Stenning (2008). I will use propositional logic for illustration purposes but the two alternative EC derivations of subtractive coercion in Chapter 3 show the same point.
The Processing of Events

Chapter 5. Processing different types of coercion 151

Involving subtractive coercion in Chapter 3. A subtractive type shift of Aktionsart will leave us with the logic program (16a) from which nothing about \( q \) can be inferred. The minimal model only contains \( p \). Second, the recomputation of the minimal model can be triggered by changing the value of \( ab \) in (16b) from undecided to true. Hence, \( \neg ab \) is false. From this it follows that \( p \land \neg ab \) is false. As a result \( q \) does not hold any more. Again a minimal model will be computed that only contains \( p \). But in this case the computation is quite different: Aktionsart is kept constant, but a defeasible inference is canceled. The latter computation corresponds to the second derivation of subtractive coercion in Chapter 3.

It is plausible that the two computations of the coerced meaning differ in processing load. Modeling both computations with recurrent neural nets lends further support for the assumption that the second computation is more difficult than the first. Adopting the framework by van Lambalgen & Stenning (2008), in Appendix A it is demonstrated that the second computation involves more steps than the first and must therefore be considered more complex.

Is it the case that subtractive coercion in this experiment was computed on the basis of a simple type shift of Aktionsart but in Experiment 1 on the basis of defeasible inference? This would explain the different patterns of results across experiments. But why should computations be qualitatively different?

Remember that the first derivation of subtractive coercion in Chapter 3 was resolved by reinterpreting the accomplishment as activity. If readers in the present study carried out reanalysis along these lines the adverbial automatically shifted the accomplishment into an activity by eliminating the \( \text{finishe} \) event. By contrast, in Experiment 1 this kind of reanalysis may be impossible because the discourse model has already been computed, that is \( \text{finishe} \) has already been made true. In Experiment 1 the incremental makes sense judgements forced participants to elaborate on a full interpretation of minimal sentences like \( \text{John built a house} \ldots \). This may lead to a similar effect as an explicit clause boundary. Again, only the accomplishment meaning may be retained. Thus, only the second computation may have been viable in these studies.

Comparing the rejection rates from Experiment 1 with the semantic acceptability judgments in this experiment provides initial evidence that goes in this direction. In Experiment 1 subtractive coercion was judged nonsensical more often than control, while in the present experiment there was no difference in acceptability. This indicates that \( \text{for-} \)adverbials had different effects in the two experiments.

Interestingly, the minimal models differ between the two computations. While a simple type shift leaves undecided whether the goal state was attained eventually, the explicit mention of a disabling condition makes the culmination false. Thus, questions which query the culmination should receive “yes” and “no” answers equally often in the first case but “no” answers in the second. The “attainment of the goal”-judgments in the present experiment pattern with the first option. This provides further support for the claim that readers computed answers on the basis of a simple activity scenario.

Two alternative ways to achieve subtractive coercion

Greatly simplifying matters by employing a non-monotonic three-valued propositional logic like the one proposed by Stenning & van Lambalgen (2005), activities and accomplishments can be represented by the logic programs given in (16). In contrast with the classical notion of validity Stenning & van Lambalgen (2005) use a non-monotonic version: “an argument is valid if the conclusion is valid in all preferred models of the premises” (p. 935).

\[(16)\]
\[
\begin{align*}
&\text{a. } \quad P_{\text{activity}} = \{p\} \\
&\quad \text{(where } p \text{ is some activity like } \text{John is running}) \\
&\text{b. } \quad P_{\text{accomplishment}} = \{p \land \neg ab \rightarrow q, p\} \\
&\quad \text{(where } p \text{ is some activity like } \text{John is building a house}, q \text{ the culminating event like the house is finished and } ab \text{ some “abnormality” preventing the culmination to happen like a credit failure happens)}
\end{align*}
\]

In (16a) the program consists of a single proposition expressing an activity. (16b) is more complex: the first formula is the logical form of an accomplishment, namely “if there is a preparatory process \( p \) and nothing abnormal (= \( ab \)) happens, then the culminating event \( q \) will happen”. Additionally, the progressive introduces the fact that there is a process \( p \). Without evidence to the contrary, closed world reasoning will make \( ab \) false. Thus the inference of \( q \) is (defeasibly) valid.

The inference of \( q \) can be canceled in two ways. First, the accomplishment can be coerced into an activity. This was also done in the first derivation of a sentence involving subtractive coercion in Chapter 3. A subtractive type shift of Aktionsart will leave us with the logic program (16a) from which nothing about \( q \) can be inferred. The minimal model only contains \( p \). Second, the recomputation of the minimal model can be triggered by changing the value of \( ab \) in (16b) from undecided to true. Hence, \( \neg ab \) is false. From this it follows that \( p \land \neg ab \) is false. As a result \( q \) does not hold any more. Again a minimal model will be computed that only contains \( p \). But in this case the computation is quite different: Aktionsart is kept constant, but a defeasible inference is canceled. The latter computation corresponds to the second derivation of subtractive coercion in Chapter 3.

It is plausible that the two computations of the coerced meaning differ in processing load. Modeling both computations with recurrent neural nets lends further support for the assumption that the second computation is more difficult than the first. Adopting the framework by van Lambalgen & Stenning (2008), in Appendix A it is demonstrated that the second computation involves more steps than the first and must therefore be considered more complex.

Is it the case that subtractive coercion in this experiment was computed on the basis of a simple type shift of Aktionsart but in Experiment 1 on the basis of defeasible inference? This would explain the different patterns of results across experiments. But why should computations be qualitatively different?

Remember that the first derivation of subtractive coercion in Chapter 3 was resolved by reinterpreting the accomplishment as activity. If readers in the present study carried out reanalysis along these lines the adverbial automatically shifted the accomplishment into an activity by eliminating the \( \text{finishe} \) event. By contrast, in Experiment 1 this kind of reanalysis may be impossible because the discourse model has already been computed, that is \( \text{finishe} \) has already been made true. In Experiment 1 the incremental makes sense judgements forced participants to elaborate on a full interpretation of minimal sentences like \( \text{John built a house} \ldots \). This may lead to a similar effect as an explicit clause boundary. Again, only the accomplishment meaning may be retained. Thus, only the second computation may have been viable in these studies.

Comparing the rejection rates from Experiment 1 with the semantic acceptability judgments in this experiment provides initial evidence that goes in this direction. In Experiment 1 subtractive coercion was judged nonsensical more often than control, while in the present experiment there was no difference in acceptability. This indicates that \( \text{for-} \)adverbials had different effects in the two experiments.

Interestingly, the minimal models differ between the two computations. While a simple type shift leaves undecided whether the goal state was attained eventually, the explicit mention of a disabling condition makes the culmination false. Thus, questions which query the culmination should receive “yes” and “no” answers equally often in the first case but “no” answers in the second. The “attainment of the goal”-judgments in the present experiment pattern with the first option. This provides further support for the claim that readers computed answers on the basis of a simple activity scenario.
To sum up, the present study tapped on a relatively easy type of coercion, namely subtractive coercion. It didn’t support the Non-Monotonic Update Hypothesis (NMUH). As outlined, this may be due to different ways how subtractive coercion can be achieved. A non-monotonic update can go unnoticed if the coercing stimulus can pick out parts of a more complex Aktionsart. How this works has been illustrated in the first derivation of subtractive coercion in Chapter 3. These considerations demonstrate that it is essential to constrain formal semantic models to delimit the space of possible computations. Only then it is possible to derive testable predictions about the computation of coerced meanings.

5.3 Abstract type shift: Experiment 6

Let’s turn to the third type of aspectual coercion – abstract type shift. Abstract type shift transforms one Aktionsart into another without changing the internal structure of the eventuality. (17) provides an example where a point in (17a) is coerced into an iterative process in (17b).

(17) a. John sneezed.
   b. John sneezed for an hour.

There is linguistic evidence that semantic types are rather flexible (Partee & Rooth 1983) and can easily be transformed from one category into another. Recently, aspectual type shift has received increased psycholinguistic interest (Piñango et al. 1999; Todorova et al. 2000a; DeVelle 2004; Traxler et al. 2002; Husband et al. 2006; Piñango et al. 2006; Brennan & Pykkänen 2007). In the psycholinguistic literature, the focus on this type of aspectual coercion has been so strong that aspectual coercion is often equated with abstract type shift. In the following, I will first introduce the notion of abstract type shift and provide some linguistic background. Then I will summarize the existing studies. The latter have already been extensively described in Chapter 2 and are repeated here once more as a reminder.

5.3.1 Natural language has flexible semantic types

Meaning comes in different types. For instance, entities (of type e) are commonly distinguished from truth values (of type t) and “higher” types are constructed on the basis of primitive types.1 Even innocent looking noun phrases (NPs) show a spectrum of types: referential NPs of type (e) (like John), predicative NPs of type (e, t) (like a dog in Fido is a dog) and quantificational NPs of type ((e, t), t) like every woman (see Partee (1986) for a discussion of flexible NP types). Linguistic evidence for flexible types comes from coordinated NPs like John and some woman. Consider the contrast in (18). He is a special type of anaphor which is in need of an e-type antecedent. In (18a), the NP John is of the right type to serve as antecedent. But the quantifier everybody in (18b) cannot serve as an antecedent showing that it is of a different type.

(18) a. John walked into the bar. He looked tired.
   b. Everybody walked into the bar. *He looked tired.

NPs of different types can be coordinated like in John and some woman (and is of type ((a, a), a)). Given their different types, how is it possible to coordinate the two NPs? Assuming flexible NP types, an explanation has been given in terms of coercion (Partee & Rooth 1983). Since the quantifier every woman is (e, t), t, John has to be lifted from type e (the constant j) to type ((e, t), t) (the generalized quantifier AP.P[?]) to fulfill the input requirements of conjunction. Intuitively, the coercing John and some woman doesn’t seem harder to understand than the non-coercing John and Mary. Does this mean that type-shifting doesn’t cause difficulty in processing?

In the aspectual domain there is also type shifting which is similar to the NP cases. For instance, drawing upon the fundamental ontological distinction between events and processes examples like John sneezed for an hour demonstrate coercion from an eventuality of type event into a process. Pulman (1997) has developed a model theoretic account of type coercion which makes use of an inventory of type-shifting operations. He extends the ontology of primitive types by adding eventualities representing states ev, processes ev, and events ev. Coercion operators are defined with respect to their input and output types, analogously to type-shifting operators in the NP domain. Pulman’s account can be translated into type-theory by introducing an operator of type ev which takes any eventuality as input and outputs a process. This operator can account for the data in (19) which illustrate that basically any Aktionsart except for states can be shifted into a process by iterating the eventuality. Because this type of coercion does not depend on the semantic structure of the Aktionsart which serves as input of the coercing operator I have called it abstract type-shift.

(19) a. For half a year, John jogged in the park.
   b. For three weeks, John played the moonshine sonata.
   c. For many years, John won the race.
   d. For five minutes, John hiccuped.

Abstract type-shifting needs more than computing iterative readings. Knowledge about the typical duration of the coerced eventuality is also needed to compute the
temporal profile of the process. If, for instance, in (19d) instead of for five minutes the adverbial for five seconds is used, the preferred interpretation consists in a single hiccup event which is stretched over a time interval of five seconds.

Coercion operators are commonly assumed to apply to a complete predication (e.g. Pulman (1997); DeSwart (1998); Rothstein (2004)). This makes the interesting prediction that the iterative reading can only be computed after both the intransitive verb and the subject phrase have been encountered. In my next experiment, a word order manipulation was used to test this prediction.

5.3.2 Abstract type shift – Difficult to perform?

Abstract type shift has recently received increased interest in psycholinguistic research. The existing studies investigated English sentences under an iterative interpretation. A short summary may be in order here, although I have already described them in more detail in Chapter 2.

The first experimental study on aspectual coercion was a cross-modal lexical decision experiment conducted by Piñango et al. (1999). Coercing (20a) versus non-coercing (20b) sentences were auditorily presented to the participants. 250ms after the offset of the coercing stimulus (for a long time) a semantically unrelated probe item (e.g. bureau) was presented on the screen for lexical decision.

(20) a. The man examined the little bundle of fur for a long time. to see if it was alive.
   b. The man kicked the little bundle of fur for a long time to see if it was alive.

Lexical decision times were reliably longer after coercing stimuli than after controls. Since less capacity was available for the secondary lexical decision task the authors take the effect as evidence for enhanced working memory load due to aspectual coercion. Unfortunately, Piñango et al. (1999) did not check plausibility across conditions. This would have been especially important because different verbs were used across sentences. In a questionnaire study DeVelle (2004) gathered plausibility ratings for the items in the two conditions. She found that control was judged more plausible than coercions. The difference in decision time in the Piñango et al. (1999) experiment may thus be an experimental artefact.

A study by Todorova et al. (2000a) on coercion used incremental makes sense judgments. In contrast to Piñango et al. (1999), Todorova et al. (2000a) kept the verb constant within item sets, as shown in the sample item in (21).

(21) a. Even though Howard sent a large check to his daughter for many years, she refused to accept the money.
   b. Even though H. sent a large check to his daughter last year,
   c. Even though H. sent large checks to his daughter for many years,
   d. Even though H. sent large checks to his daughter last year.

They found longer reading times and increased rejection rates at the adverbial region of (21a) compared to (b–d). Note that (21a) is the only condition which triggers an iterative reading since a punctual verb is combined with a durative for-adverbial. In (21c) a plural object is used which makes the VP durative and in (21b) and (21d) a punctual adverbial serves as control. Thus, the findings seem to suggest that computing an iterative reading is more difficult than the processing of non-iterating controls, at least in an offline task like incremental makes sense judgments.

However, while the items included punctual verbs, there was no discrimination between semelfactives (e.g. kicked, hit) and achievement verbs (e.g. entered, found). Moreover, the use of definite rather than indefinite articles was ungrammatical with achievement verbs: although the dragon devoured a girl from the village for years vs. although the dragon devoured the girl from the village for years. Such variations make it difficult to ascertain the source of slower reading times. What is badly needed over and above the reported results are grammaticality judgments of the conditions in (21).

In a series of two self-paced reading and two eyetracking studies Pickering et al. (2006) tried to replicate the studies of Piñango et al. (1999) and Todorova et al. (2000a). They employed an ordinary reading task since both earlier studies used methods that may not reflect normal language comprehension. Interestingly, in none of their experiments did they find processing difficulty due to coercion. The authors take this as evidence for the aspectual underspecification hypothesis which states that during ordinary reading the Aktionsart remains underspecified (at least until the the end of the sentence). However, explicitly querying the reading was never used in their experiments to test whether subjects actually computed iterative readings. Thus, an alternative explanation for the results is that readers only did shallow processing.

Recently, Piñango et al. (2006) replicated the results of their earlier study, again using the cross-modal lexical decision task. Since the items were similar to those of the first experiment the same criticism can be applied to this study, too. In their Experiment 2 they manipulated the time between offset of the coercing stimulus and the appearance of the probe stimulus. While lexical decision times were different in the coercing versus control condition in Experiment 1 with an asynchrony of 250 ms, no difference was detectable with synchronous presentation. If the findings of Experiment 1 do in fact reflect difficulty caused by aspectual coercion, this result is interesting since it suggests that aspectual composition is a rather late process.

Finally, the contributions by Husband et al. (2006) and Brennan & Pylkkänen (2008a) provide self-paced reading and MEG data on aspectual coercion. The first study reports two self-paced reading experiments. Experiment 1 replicated the results by Todorova et al. (2000a). Punctual VPs combined with a for-adverbial took longer to read than controls. Since the same materials were used as in Todorova et al. (2000a) the same criticism can be applied.
The study by Brennan & Pykkänen (2008a) investigated the computation of iterative readings using self-paced reading and MEG in constructions with unambiguous punctual verbs. The self-paced reading experiment resulted in an effect with coercion being read more slowly than control. The MEG data gave same indication of prefrontal activation in the coercion condition (Throughout the day the student sneezed ...) compared to aspectual control (After twenty minutes the student sneezed ...). The authors point out that the effect resembles the anterior midline MEG effect elicited by complement coercion (Pykkänen, L. & McElree, B. 2007). In addition, preceding the anterior midline field they found activation resembling the brain response to semantic anomaly. They conclude that aspectual coercion first leads to mismatch detection which is followed by semantic enrichment which is carried out in prefrontal areas of the brain.

To sum up, the computation of iterative readings has been reported to be costly in cross-modal lexical decision and in the stops-making sense task but studies employing more natural reading have yielded mixed results. So far, all studies have investigated English sentences. Since the temporal concepts underlying Aktionsart seem to be universal (Dahl 1985; Smith 1991), it is interesting to study the processing of iterative readings in a different language, too. Moreover, in most of the studies there were problems concerning the materials. To face these problems it is necessary to carefully control factors like plausibility, grammaticality, verb type, frequency etc. across conditions. Finally and most crucially, in order to be able to interpret null effects one has to determine whether comprehenders really computed iterative readings. The following self-paced reading experiment was aimed at overcoming the problems mentioned before.

5.3.3 The abstract type shift hypothesis (ATSH)

The Abstract Type Shift Hypothesis from Chapter 3 is repeated in (22).

(22) **Abstract Type Shift Hypothesis (ATSH)**

Abstract type shifting causes processing difficulty.

The ATSH predicts that transformations of an Aktionsart induce difficulty regardless of whether there is any modification of its internal structure. Hence, in experiments studying the online processing of iterated readings we expect to find processing difficulty in coercing an event into an iterated activity reading.

Further, in some semantic theories of abstract type-shifting it is assumed that type-shift is achieved by the application of a covert operator which transforms a whole predicate-argument structure. This leads to the interesting prediction that abstract type-shift should only be possible when readers have accessed the complete verb-argument structure. Thus, we expect a mediating influence of word order on aspectual coercion effects.

The ATSH makes the following predictions with regard to the constructions in (23). In (23a) the for-adverbial shifts the semelfactive a student sneezed into an iterated process. (23b) does not involve iteration. Hence the adverbial phrase in (23a) should be more difficult to process than in (23b). (23c) and (23d) are word-order variants of (23a) and (23b). Again, the ATSH predicts (23c) to be more difficult than (23d). However, if it is the whole verb-argument structure that undergoes abstract type shift, difficulty should emerge at the earliest when the subject eine Studentin is encountered. In contrast, if the bare verb has the potential to express semelfactive Aktionsart on its own, difficulty should already emerge at the adverbial phrase.

(23) a. Eine Studentin nieste mehrere Tage überaus laut auf der Exkursion…
   At the field trip, a female student sneezed rather loudly for several days…

b. Eine Studentin nieste gerade eben überaus laut auf der Exkursion…
   At the field trip, a female student sneezed just now rather loudly…

c. Auf der Exkursion nieste mehrere Tage überaus laut eine Studentin,…
   At the field trip sneezed for several days rather loudly a female student…

d. Auf der Exkursion nieste gerade eben überaus laut eine Studentin,…
   At the field trip sneezed just now rather loudly a female student…

2. The interpretation is somewhat problematic since they compare the computation of an iterative reading to a semantic anomaly that has nothing to do with aspect (their comparison is Halgren et al. (2002)). It would have been nice if they included an aspeuctal mismatch condition in their experiment.
5.3.4 Method

The ATSH was tested in Experiment 6 using self-paced reading with moving-window presentation. To avoid the problems raised in connection with the above-mentioned studies, two pretests were conducted which controlled for plausibility and grammaticality. In addition, to make sure that the extraposed subjects could not be predicted on the basis of the verb alone, a sentence completion pretest was run. Finally, to guarantee that readers actually computed abstract type shift, half of the experimental trials were followed by questions which queried iterative readings.

Materials

40 candidate items were constructed according to the scheme in (23). Items were segmented in a way that always at least two spill-over regions followed both the critical adverbial and the subject phrase. In (24) slashes indicate segmentation.


The whole day, the female student sneezed during the field trip very loudly, because she herself a cold gotten had.

Whether a verb is represented as punctual might depend on the likelihood that the action described by the verb is naturally repeating. Only such German semelfactive verbs were chosen that most naturally describe a single event. The 40 potential items were subjected to the pretests described in the next section. On the basis of the pretest results 24 experimental items were chosen. A list of sentences is provided in Appendix I.

Items were controlled for length and frequency of the adverbial phrase, the only phrase that differed between coercion and control conditions. There was neither a difference in length (mean length of coercing adverbials: 15.5 characters, control: 15.8 characters; results from an independent samples t-test: \( t(23) = .26; p = .799 \)) nor of frequency (mean absolute frequency of coercing adverbials: \( \text{freq} = 1777 \), control: \( \text{freq} = 1596 \); results from an independent samples t-test: \( t(23) = .29; p = .771 \)). Frequency was normalized and then subjected to \( 2 \times 2 \) repeated measures ANOVAs (factors: coercion and word order). ANOVAs including only the 24 items selected for the experiment did not reveal an effect either of coercion \( (F_{1/2} < 1) \), or of word order \( (F_{1, 19} = 1.39; p = .26; F_2 < 1) \) or their interaction \( (F_{1/2} < 1) \). Thus, the selected items are equally plausible in all conditions.

- plausibility: In connection with the studies by Piñango et al. (1999) and Piñango et al. (2006) the concern was raised that differences in lexical decision times might reflect unequal plausibility among conditions instead of enhanced processing cost due to abstract type shift. To exclude this possibility in the present experiment, a norming study was conducted over the internet (using WebExp2 software, Mayo et al. (2006)). 20 participants provided ratings for the 40 items (+ 60 fillers) on a seven point scale with randomized order of presentation. Items and fillers were assigned to four lists according to a Latin square design. The judgments were normalized and then subjected to \( 2 \times 2 \) repeated measures ANOVAs (factors: coercion and word order). ANOVAs including only the 24 items selected for the experiment did not reveal an effect either of coercion \( (F_{1/2} < 1) \), or of word order \( (F_{1, 19} = 1.39; p = .26; F_2 < 1) \) or their interaction \( (F_{1/2} < 1) \). Thus, the selected items are equally plausible in all conditions.

- grammaticality: The studies by Todorova et al. (2000a) and Husband et al. (2006, Experiment 1) had the problem that in the aspectual coercion condition some items were grammatically marked. For this reason, in the present experiment sentences were checked for grammaticality. Further, since Experiment 6 investigated extraposed subjects I wanted to make sure that this construction type was not perceived as marked. A grammaticality test of the candidate sentences was performed by gathering grammaticality judgements from 20 participants using the magnitude estimation method (Bard et al. (1996). This allowed us to compare the sentences with 40 normed distractors which included a whole spectrum of grammatical violations showing different degrees of markedness (sorted in five categories from “perfect” to “strongly marked” with 4 fillers in each category). The list of fillers is contained in Appendix I. Four lists were constructed according to a Latin square design. The questionnaire was administered over the internet using WebExp2 software. Figure 5.6 shows the judgments for the 24 experimental items. Although the coercion conditions were judged as slightly more marked than the controls, in terms of absolute grammaticality all conditions were roughly the same (category B which contained complex but fully grammatical constructions). Further, extraposed subjects were judged as grammatical as the canonical order.

Pretests

The 40 candidate items were tested in three pretests. Based on the pretest results 24 items were selected for the reading time study. The results are reported only with respect to the 24 items which were included in the experiment.

3. Thanks to Sam Featherston for kindly providing the normed materials which were used as distractors in the questionnaire.
The Processing of Events

Chapter 5. Processing different types of coercion

Participants were asked to finish the sentences in a meaningful way. The 40 items were distributed over four lists according to a latin square design. A continuation was judged as matching the subject phrase if the same concept, a synonym, a specialized term or a superordinate term was used (e.g. if the target concept was *hand gun, pistol, Magnum 44* or *gun* would all count as hits). The percent of hits of the 24 experimental items was computed and subjected to statistical analysis. In the short conditions there were only 3.9% hits in the coercion condition (25a) and 8.6% hits in the control condition (25b). The long conditions had 20.8% in (25c) and 19.9% in (25d) respectively. ANOVAs with the two factors *coercion* (*coercion vs. control*) and *length* (*short vs. long*) revealed a significant main effect of length ($F_1(1, 23) = 30.16; p < .01; F_2(1, 23) = 11.27; p < .01$). Neither the main effect of *coercion* ($F_{1/2} < 1$) nor the interaction ($F_1(1, 23) = 1.82; p = .19; F_2 < 1$) was significant. Not surprisingly, participants were more successful in guessing the subject when more sentence information was available to them than when the sentence fragment was rather short. Most crucially, however, is that the subject phrase was not predictable at the segment containing the adverbial phrase.

Participants

24 native German speakers recruited at Tübingen University participated in the experiment (14 female, mean age = 25.1, ranging from 20 to 44). All participants were naive to the purpose of the study. Each subject was paid 5 Euros for participation. The participants were randomly assigned to lists (six subjects per list).

Procedure and data analysis

The procedure was identical to the previous experiments. Reading time data and question answering data were analyzed like in the previous experiments. An experimental session lasted approximately 30 minutes.

5.3.5 Results

In the following, I will first report question answering data and then turn to reading times. Questions querying iterative readings, henceforth *iteration questions*, followed half of the items ($n = 12$) to check whether the experimental manipulation was successful. They were analyzed separately from ordinary content questions.

Question answering data

Averaged across all trials (experimental items and fillers), content questions were answered correctly 91.4% of the time. ANOVAs revealed that performance on content questions was equal across conditions (all $F_{1/2} < 1$) with a mean of 90.6% correct answers. Thus, participants read attentively.

---

The cloze probabilities: A potential early effect on the adverbial phrase with an exposed subject can only be interpreted if participants cannot predict the lexical content of the yet missing subject phrase on the basis of the preceding verb. To rule out this possibility, a sentence completion pretest was run (again over the internet using WebExp2). 24 participants provided continuations for sentence fragments of varying length. Four conditions were tested:

(25) a. Auf der Exkursion nieste den ganzen Tag
   At the field trip sneezed for the whole day
b. Auf der Exkursion nieste gestern früh
   At the field trip sneezed yesterday morning
c. Auf der Exkursion nieste den ganzen Tag überaus laut
   At the field trip sneezed for the whole day very loudly
d. Auf der Exkursion nieste gestern früh überaus laut
   At the field trip sneezed yesterday morning very loudly

Participants were asked to finish the sentences in a meaningful way. The 40 items were distributed over four lists according to a latin square design. A continuation was judged as matching the subject phrase if the same concept, a synonym, a specialized term or a superordinate term was used (e.g. if the target concept was *hand gun, pistol, Magnum 44* or *gun* would all count as hits). The percent of hits of the 24 experimental items was computed and subjected to statistical analysis. In the short conditions there were only 3.9% hits in the coercion condition (25a) and 8.6% hits in the control condition (25b). The long conditions had 20.8% in (25c) and 19.9% in (25d) respectively. ANOVAs with the two factors *coercion* (*coercion vs. control*) and *length* (*short vs. long*) revealed a significant main effect of length ($F_1(1, 23) = 30.16; p < .01; F_2(1, 23) = 11.27; p < .01$). Neither the main effect of *coercion* ($F_{1/2} < 1$) nor the interaction ($F_1(1, 23) = 1.82; p = .19; F_2 < 1$) was significant. Not surprisingly, participants were more successful in guessing the subject when more sentence information was available to them than when the sentence fragment was rather short. Most crucially, however, is that the subject phrase was not predictable at the segment containing the adverbial phrase.
The Processing of Events

Figure 5.7. Conditions: no significant main effect of coercion (SD = 16.8) in the vas-control. Consequently, ANOVAs yielded no differences between coercion condition with vas-order versus 53.9ms/char. (SD = 16.2) in the sva-control and were numerically even faster than the controls: 50.0ms/char. (SD = 17.5) in the coercion condition with vas-order versus 51.6ms/char. (SD = 13.3) in the sva-order versus 51.0ms/char. (SD = 21.7; p < .01). There was no significant main effect of word order (F_{1,2} < 1) and no interaction between word order and coercion (F_{1,2} < 1).

Question answering times didn’t differ across conditions, either with respect to content questions or to iteration questions (all p – values > .1).

Reading times

Figure 5.7 depicts the reading times. At the critical adverbial the coercion conditions were numerically even faster than the controls: 50.0ms/char. (SD = 17.5) in the coercion condition with sva-order versus 53.9ms/char. (SD = 16.2) in the sva-control and 51.0ms/char. (SD = 13.3) in the coercion condition with vas-order versus 51.6ms/char. (SD = 16.8) in the vas-control. Consequently, ANOVAs yielded no differences between conditions: no significant main effect of coercion (F_{1,2} = 3.717; MSE = 121.91; p = .19; F_{1, 2} = 1.336; MSE = 100.41; p = .19), no main effect of adverbial (all F_{1,2} < 1) and no interaction (all F_{1,2} < 1). At the following regions, the two word orders were analyzed separately by paired t-tests because the segments differed lexically. Reading times did not differ either at the spillover region of the adverbial (both t_{1/2} < 1), or at the following segment which contained the extraposed subject phrase in the vas-order (both t_{1/2} < 1) or the following region (both t_{1/2} < 1).

To further investigate whether coercion might have taken place rather late in the sentence I computed end of sentence reading times starting with the adverbial. These are shown in Table 5.2.

Table 5.2. End of sentence reading times in Experiment 6

<table>
<thead>
<tr>
<th></th>
<th>Mean (Mean standard error)</th>
</tr>
</thead>
<tbody>
<tr>
<td>sva order</td>
<td></td>
</tr>
<tr>
<td>for x time</td>
<td>5501 (131)</td>
</tr>
<tr>
<td>at x time</td>
<td>5492 (131)</td>
</tr>
<tr>
<td>vas order</td>
<td></td>
</tr>
<tr>
<td>for x time</td>
<td>5245 (133)</td>
</tr>
<tr>
<td>at x time</td>
<td>5132 (123)</td>
</tr>
</tbody>
</table>

Note: end of sentence reading times (in ms) have been computed by summing up raw reading times (corrected for outliers) over regions.

ANOVA revealed a significant main effect of word order in the analysis by participants (F_{1,23} = 11.35; MSE = 2269708.76; p < .01; F_{1, 23} = 2.83; MSE = 2269708.76; p = .11) which reflects faster reading times in vas sentence endings (which were a few syllables shorter for some items) compared to sva endings. For x time adverbials, however, didn’t differ from at x time-adverbials and there was no interaction between word order and adverbial (all F_{1/2} < 1).

5.3.6 Discussion

Experiment 6 tested whether abstract type-shift causes processing difficulty as stated in the ATSH. The study used materials that were carefully controlled with respect to potential confounds such as plausibility across conditions. The answers to iteration questions clearly indicate that the combination of a punctual semelfactive with a for-adverbial leads to an iterative reading. The reading time data, however, do not support the ATSH, since computing an iterative reading did not cause any processing difficulty. On the contrary, the coercing adverbials were numerically even faster than the non-coercing controls.

The results of the present experiment are in line with the study by Pickering et al. (2006) who also found no difficulty in abstract type-shift. They take the null effects in reading times as evidence for their aspectual underspecification theory. By contrast, in my opinion the findings only indicate that abstract type-shift is a particularly easy kind of aspectual coercion. The aspectual underspecification hypothesis cannot be true because additive coercion in Experiment 4a lead to a considerable amount of difficulty during ordinary reading.
Although the interpretation of null-effects is always problematic the present findings are incompatible with those by Piñango et al. (1999), Piñango et al. (2006), Todorova et al. (2000a), Brennan & Pylkkänen (2008a) and Husband et al. (2006). The discrepancy might be due to the above-mentioned confounds. An alternative explanation, however, might be that abstract type-shifts are unequally difficult across languages. Why should that be?

Nouns are often derived from verbs via imperfect nominalization (Vendler 1968). (26) is an example. In this type of nominalization, an event has to be given a fluent interpretation. It involves the same semantic operation that is required when computing iterative readings.

(26) Peters (dauernes) Niesen … Peter’s (perpetual) sneezing …

A quick corpus research querying all written corpora of the Cosmas corpus showed that the verbs used in the present experiment appear rather often in their nominalized form. Comparing the absolute occurrences of the verb lemmas either beginning with a small letter (e.g. niesen; to sneeze) or beginning with a capital letter (e.g. Niesen; the sneezing) revealed that twelve out of twenty four verbs that were used in Experiment 6 occurred more often in a nominalized form than in a verbal form. This indicates that in German, transformations from semelfactives into process denoting nouns are extremely common. This provides an explanation why process readings may be easily accessible. In English, however, the situation may be entirely different. With respect to the ATSH it would be particularly interesting to study languages which don’t have nominalizations at all (Koptjevskaja-Tamm (2005) lists 42 such languages in a sample of a total of 168 languages). The computation of iterative readings may be cognitively more demanding in these languages since abstract type-shift is less common.

To sum up, language makes extensive use of flexible type-shifting. Often, a type-conflict is repaired without even noticing it. In line with this observation, the present study provides evidence that abstract type-shift does not induce measurable cost during language processing. The findings of the present experiment were replicated in Experiment 10b lending additional substance to the claim that this type of coercion doesn’t cause any difficulty. The ATSH is thus disconfirmed by the experimental results.

4. But since Experiment 10b in Chapter 8 replicates the results of the present experiment, the findings of the present study seem to be reliable.

5.4 General discussion and conclusions

In this chapter I have investigated processing difficulty in three different kinds of aspectual coercion. I used a classification system which has partly been proposed by Hamm & van Lambalgen (2005).

Aspectual coercion is additive if the aspectual representation has to be enriched by adding an eventuality to it. Additive coercion was investigated in Experiment 4a using achievements which were coerced into accomplishment. The experiment provides evidence for enhanced processing difficulty in additive coercion.

Subtractive coercion is the negative counterpart to additive coercion. This kind of coercion forces the processor to cancel some part of the eventuality. Experiment 5, which investigated the transformation of an accomplishment into an activity, didn’t find difficulty in subtractive coercion. But since subtractive coercion may be achieved in different ways issues aren’t settled yet and further research is required.

Finally, in case of abstract type-shifting the aspectual representation may be coerced as a whole without modifying the “embedded” eventuality. The computation of iterative readings provides an instance of flexible type-shift. This type of coercion has already received a fair amount of psycholinguistic investigation without conclusive evidence yet. Experiment 6 investigated the computation of iterative readings in German semelfactives which were modified by a durative adverbal. There was no indication of difficulty in this experiment.

The three kinds of coercion rest upon three notions which play a key role in natural language semantics. Additive coercion rests upon context-sensitivity and can only be achieved when sentential and contextual information are both taken into account. Subtractive coercion violates monotonicity since an inference that was valid before might be invalid after the coercing stimulus has been interpreted. Abstract type shift is based upon a flexible type system which allows to move from one type to another.

What is striking about the findings reported in this chapter is not so much that difficulty was found in additive coercion but rather that non-monotonic update and abstract type-shift seem to cause no difficulty. This indicates that the latter operations may be so common in language comprehension that they can be performed completely unnoticed.
The processing domain of Aktionsart

In the semantic literature on Aktionsart it is commonly assumed that Aktionsart is a property of VPs or even of whole sentences. Does the interpretation of aspect have to wait until the verb plus all its arguments are present? To address this issue, two self-paced reading studies tested German examples of additive coercion like (1a) where an achievement is coerced into an accomplishment, as well as controls like (1b) and aspectual mismatch like (1c). The latter were included as other coercion phenomena have been reported to show quite late effects (e.g. Traxler et al. (2002)).

(1) a. Hans erreichte den Gipfel in zwei Tagen von der Nordseite her.
   Hans reached the top in two days from the northface.

   b. Hans erreichte den Gipfel vor zwei Tagen von der Nordseite her.
   Hans reached the top two days ago from the northface.

   c. Hans erreichte den Gipfel zwei Tage lang von der Nordseite her.
   Hans reached the top for two days from the northface.

I tested aspectual processing at a point where the argument structure of a clearly transitive verb is not complete yet. However, readers may predict arguments that shift the intended Aktionsart (e.g. a bare plural in John won races). To exclude this possibility a combined acceptability rating/completion study was conducted (Experiment 7). The participants judged if the initial words of the items (like Zwei Tage lang erreichte …) could be continued in a meaningful way and if so provided a continuation. Three conditions were tested: adverbial-verb, object-verb-adverbial or subject-verb-adverbial and always used for-adverbials which force subjects to choose arguments that shift the Aktionsart. Across conditions, the fragments were overwhelmingly judged as nonsensical indicating an achievement interpretation.

Experiment 4 tested whether Aktionsart is already determined at the verb. (1a–c) were compared in two orders: the canonical subject-verb-object-adverbial order (see Experiment 4a in the preceding chapter) versus subject-verb-adverbial-object (Experiment 4b). While in Experiment 4a both an aspectual mismatch effect and a coercion effect occurred at the adverbial, in Experiment 4b there was no indication of difficulty. This lack of effect in the verb-adverbial-object order suggests that the parser does not determine Aktionsart on the basis of the verb alone.
The increment size in aspectual processing

In theories of language processing it is commonly assumed that syntactic interpretation proceeds incrementally, that is on a word by word basis. An open question is whether this holds for aspectual semantic processing and for semantic processing in general as well. Crocker (1996, p. 251) formulated the principle of incrementality (the psycholinguistic perspective on syntactic processing) in the following way:

“The sentence processor operates in such a way as to maximize the interpretation and comprehension of the sentence at each stage of processing (i.e. as each lexical item is encountered).”

By contrast, in semantic theory Aktionsart is often treated as a property of whole VPs or even whole sentences. This is what I call the semantic perspective (Dowty 1979, p. 62):

“Not just verbs, but in fact whole verb phrases must be taken into account to distinguish activities from accomplishments. (In a certain sense, even whole sentences are involved...).”

If we take the semantic perspective seriously, we expect that a transitive verb on its own has no Aktionsart until it is composed with (at least) its internal argument. As a consequence, effects due to aspectual violations and aspectual coercion can only arise when the verb has received all or at least some of its arguments.

Consider the examples in (2a–c). Note that erreichen (reach) is an unambiguous transitive achievement verb. Hence it doesn't allow modification by a for-adverbial rendering all three word order variants ungrammatical.

(2) a. Der Bergsteiger erreichte den Gipfel *zwei Stunden lang. the mountaineer reached the top for two hours
   b. Den Gipfel dir.obj erreichte *zwei Stunden lang ein Bergsteiger subj. The top reached for two hours a mountaineer
   c. Der Bergsteiger subj erreichte *zwei Stunden lang den Gipfel dir.obj. The mountaineer reached for two hours the top

What makes the three examples interesting is when the aspectually mismatching information comes into play: In (2a) the verb-argument-structure is complete when the adverbial enters the sentence. In (2b) the verb has already received the direct object, but the subject is still missing. This means we are dealing with a complete VP. Finally, in (2c) the VP is actually not complete yet. At this point, the adverbial has to modify the bare verb.

The present chapter investigates whether an aspectual violation can be detected immediately at the mismatching adverbial irrespective of its structural position in the sentence. Further, since coercion may require more contextual information than mere mismatch detection, even more deferred processing may be expected in coercion than in mismatch. The time-course of aspectual violation and reanalysis will be investigated with word order variants of the construction used in Experiment 4a of the preceding chapter. This study established that additive coercion leads to a considerable amount of processing difficulty. Interestingly, mismatch and coercion slowed down reading of the adverbial phrase way before the sentence was over. Accordingly, the starting point of the present chapter is that at least a minimal sentence is aspectually specified.

6.1.1 Can the existing studies tell us anything about the domain size?

With the exception of Baggio et al. (2008) all existing studies on aspectual coercion have used English materials. Because English has fixed word order it cannot be used to systematically investigate the processing of Aktionsart at various hierarchical levels.

For instance, to test the VP as processing domain, the most natural choice is to use a transitive verb in a sentence with object before subject word order where the coercing stimulus intervenes between the VP (= verb + direct object) and the subject. Unfortunately, this word order is ungrammatical in English. Thus, a language with relatively free word order like German is needed where all three construction types in (2a–c) are grammatical.

Not surprisingly, the processing domain of Aktionsart has not been explicitly mentioned in the psycholinguistic literature. Let’s have a look at the materials used in these studies to see if there is any implicit evidence concerning the issue. The following examples present sample materials from the studies reporting a coercion effect:

(3) a. The insect glided effortlessly until ...
   b. The insect hopped effortlessly until ...
   c. Howard sent a large check to his daughter for many years ...
   d. Howard sent a large check to his daughter last year ...

Sentences like (3) were used in the studies by Piñango et al. (1999) and Piñango et al. (2006). The coercing adverbial (until ...) only appeared after a minimal sentence was complete. Similarly, the materials in (4) used by Todorova et al. (2000a) and by Husband et al. (2006) only allowed to detect a coercion effect after a complete verb-argument structure had been presented. Finally, the study by Baggio et al. (2008) was interested in non-monotonic updates of the discourse representation which were triggered by the verb in sentence-final position. Taken together, all online effects that have been reported were measured rather late downstream the sentence. Therefore,
The existing studies do not let us decide between the incremental aspectual interpretation hypothesis in (5) and the late aspectual interpretation hypothesis in (6).

(5) **Incremental Aspectual Interpretation Hypothesis (IAIH)**
Aktionsart is computed incrementally, on a word-by-word basis.

(6) **Late Aspectual Interpretation Hypothesis (LAIH)**
Aktionsart is not computed before a verb has all its arguments.

The IAIH and its counterpart, the LAIH, are the two extremes with respect to incrementality. Certainly, there is also an intermediate alternative. Not the complete sentence, but only the VP may constitute the processing domain of Aktionsart. This last option is the complete verb phrase hypothesis (CVPH) of Chapter 3.

6.1.2 Pretesting the grammaticality of the word order variants

Although German exhibits a relatively free word order it is not entirely free. This means that some word order variants are more marked than others. This is illustrated by the extensively studied subject-object ambiguities in German which show graded effects of markedness (e.g. Bader & Meng (1999)). To compare aspectual processing among different syntactic configurations it is crucial that the constructions under study do not differ in grammaticality.

**Method and results**

Grammaticality judgments for word order variants of the items used in the online experiments were gathered using the thermometer judgement method (Featherston 2008). The following orders were tested: subject-verb-object-adverbial (svoa), subject-verb-adverbial-object (svao) and object-verb-adverbial-subject (ovas).\(^1\) All sentences were semantically well formed and used a transitive achievement modified by a x time ago-adverbial (this combination was the control condition in the reading time studies). Here is a sample item:

svoa) Der Bergsteiger erreichte den Gipfel vor zwei Stunden.
The mountaineer reached the top two hours ago.

svao) Der Bergsteiger erreichte vor zwei Stunden den Gipfel.
The mountaineer reached the top two hours ago the top.

ovas) Den Gipfel erreichte vor zwei Stunden der Bergsteiger.
The top reached two hours ago the mountaineer.

Besides relative differences I was also interested in absolute degrees of grammaticality to find out if all three constructions are perceived as fully grammatical. To this end, 20 normed distractors of five different levels of grammaticality were included.\(^2\) Categories A and B included perfectly grammatical sentences with sentences of category B being slightly more complex and stylistically not as good as sentences of category A. Categories C (= mildly marked) and D (= strongly marked) contained marked constructions and category E encompassed ungrammatical sentences. A list of the normed distractors is provided in Appendix L.

The 30 experimental items and the distractors (20 normed fillers + 20 fillers from an unrelated experiment) were used to construct four lists according to a latin square design. 20 German native speakers rated the sentences in a questionnaire study which was administered over the internet using WebExp2 software (Mayo et al. 2006). For each participant, the ratings were normalized by computing z scores on the basis of all the judgments he had provided. Figure 6.1 shows the results of the pretest. The horizontal line indicates the border between fully grammatical (cat. A and B) and marked constructions (cat. C to E).

![Figure 6.1 Mean grammaticality judgments of the three word order variants (+95% confidence intervals)](image)

1. Additionally, I also tested adverbial-verb-subject-object (avso), but since this variant wasn’t used in the online experiments I will not report the results.

2. These were chosen from a pool of German example sentences which have been repeatedly tested in grammaticality surveys at the SFB 441, e.g. Featherston (2008).
The experimental items in all three conditions were judged roughly at the same level as the grammatical fillers of category B. Paired t-tests revealed that the svoa order \( (t,(19) = −.25; p = .80) \) and the ova word order \( (t,(19) = −1.69; p = .11) \) were not reliably different from the B fillers. Interestingly, the svao order was even judged better than fillers of category B \( (t,(19) = 2.37; p < .05) \). Taken together, all three word order variants are perceived as fully grammatical.

There were, however, relative differences among the three construction types. The svao order was judged better than the svoa order \( (t,(19) = 3.94; p < .01; t,(29) = 4.65; p < .01) \) and the svao order was judged better than extraposed subjects in the ova order \( (t,(19) = 2.39; p < .05; t,(29) = 6.3; p < .01) \). Thus, in the reading time experiments an adverbial intervening between verb and object should even be easier than adverbials following the object. Sentences with extraposed subjects are slightly more marked than sentences with a svoa order. To compensate this difference, the ova construction will be tested in an experiment that only uses object initial sentences in the items and in the fillers.

6.2 What do readers predict? A production experiment (Exp. 7)

Investigating adverbial modification in yet incomplete verb–argument structures raises an important question. Do readers automatically predict an argument that yields the Aktionsart which is required by the adverbial? Consider (7) with the two continuations in (7a) and (7b).

(7) Der Bergsteiger erreichte zwei Stunden lang …
   The mountaineer reached for two hours …
   a. *den Gipfel
      the top
   b. niemanden am Telefon
      nobody on the phone

As (7b) shows, (7) can be continued in a meaningful way, although the most typical continuation of an yet incomplete achievement like the mountaineer reached … in (7a) yields a semantically ill-formed sentence. When the processor encounters the sentence fragment in (7), it will predict material that is yet to come. Let’s assume that the IAHI is correct. Then the predictive capabilities of the parser are absolutely crucial and lead to different expectations about when processing difficulty emerges in sentences like (7) with the semantically anomalous continuation (7a). There are two theoretical options concerning the predictive power of the parser. The first consists in the “charitable” parser hypothesis in (8).

(8) “Charitable” Parser
The parser predicts a sentence completion that makes the sentence sensible (if there is one at all).

A “charitable” parser will interpret the incomplete sentence in (7) with the expectation of a continuation like (7b). As a result, including the adverbial, the sentence fragment will be predicted to be well formed. Only when a continuation like (7a) is encountered, the expectation is disconfirmed and processing difficulty emerges. Thus, the charitable parser hypothesis predicts delayed processing due to incremental interpretation with extremely high predictive power.

The second option consists in the “context dependent” parser in (9).

(9) “Context Dependent” Parser
The parser expects a continuation that is highly associated with the lexical material encountered so far.

Interestingly, although the context dependent parser has less predictive power than the charitable parser, it predicts earlier difficulty in aspectual processing. In the context of the mountaineer reached something like the top is expected. The predicted object is semantically incongruous with the for-adverbial. Thus, difficulty is expected immediately at the adverbial even before the object is encountered.

Suppose in an online experiment we find delayed difficulty, that is only after the verb has received all its arguments. This hypothetical result is consistent with both late and incremental aspectual processing. It is obviously consistent with the IAHI, but the IAHI also predicts late effects if the processor is engaged in charitable prediction. We are in a dilemma. The online results can’t tell us which of the alternatives is correct. Instead, we need to get hands on the interpretation of the incomplete sentence in (7) and have to find out what readers actually predict what is yet to come.

To this end, I conducted a combined acceptability rating/sentence completion study in which readers had to judge whether the initial words of the experimental sentences could be continued in a meaningful way and if so had to provide a continuation. While charitable parsing predicts that readers will easily come up with a meaningful continuation, context dependent parsing makes the opposite prediction, namely that readers will judge the sentence fragments not to allow for a meaningful continuation.

6.2.1 Method
Experiment 7 was a production experiment with no time pressure. This ensures that participants have the opportunity to find the most sensible continuation. If readers engage in context-dependent parsing in an offline task like this it is even more probable that they will do so during real-time comprehension.
Materials
The thirty items from Experiments 4a/b and Experiment 8 were tested in the aspectual mismatch condition: an achievement combined with a for-adverbial. I eliminated the ends of the experimental sentences. This yielded the conditions in (10).

(10) a. Der Bergsteiger
erreichte
zwei Stunden lang …
   The mountaineer reached for two hours …
b. Den Gipfel
erreichte
zwei Stunden lang …
   The top reached for two hours …
c. Zwei Stunden lang erreichte …
   For two hours reached …

(10a) contains the case disambiguated subject der Bergsteiger, an unambiguously transitive achievement verb and a for-adverbial, but the object is still missing. In (10b) the case-disambiguated object den Gipfel is realized preverbally in so called Vorfeld position, but the subject is missing. In (10c) the bare verb is tested with the adverbial. In this condition readers have maximal freedom in choosing the appropriate arguments to satisfy the input requirements of the adverbial.

In the semantic literature, the prototypical examples to demonstrate that semantic properties of the arguments are important for the Aktionsart involve bare plural arguments (e.g. Verkuyl (1993)). This raises the interesting question whether readers will choose bare plural subjects more often if the numerus of the verb requires a plural subject than if it is singular like in (10a–c). To test this, I also manipulated the numerus of the verb in a factorial design with the factors verb argument structure and numerus yielding a total of six conditions; three singular in (10a–c) and three plural in (11a–c).

(11) a. Die Bergsteiger
erreichten
zwei Stunden lang …
   The mountaineers reached for two hours …
b. Den Gipfel
erreichten
zwei Stunden lang …
   The top reached for two hours …
c. Zwei Stunden lang erreichten …
   For two hours reached …

Additionally, 40 distractors were included in the experiment. 30 of them allowed for a sensible continuation while 10 clearly did not. The latter contained tense violations like morgen kam … (tomorrow came…) and aspectual violations of a different sort such as Hans war gerade dabei intelligent zu sein, als … (Hans was being intelligent, when …). The experimental items and the filler sentences were arranged in six lists according to a latin square design.

Procedure
The experiment employed a combined acceptability rating/sentence completion task. Subjects were asked to come up with a meaningful completion of the sentence. If they were not able to do so, they were prompted to reject the sentence as nonsensical.

The experiment was administered over the internet using a questionnaire that was developed with WebExp2 (Mayo et al. 2006). Only the initial part of the sentence appeared on the screen. Beneath the sentence a text field was presented in which the continuation could be typed in. Besides the text field there was a “nonsense” button. After typing in a continuation or pressing the button the next trial started.

First, subjects read written instructions. Then they accomplished a practice with five trials. After that, they received the experiment in one block with randomized presentation of 70 sentence beginnings. An experimental session lasted no longer than half an hour.

Participants
60 German native speakers (23 female; mean age 29.4 years, min. 19, max. 56) took part in the experiment. Among the participants, six prices of 50 euros were raffled in a lottery. Participants were randomly assigned to lists (10 subjects per list).

Data analysis
Data were analyzed both quantitatively and qualitatively. For purposes of quantitative analysis, the percent of “nonsense” ratings were computed. In addition to “nonsense” button presses, all continuations which yielded sentences which were not sensible or incomplete were also counted as “nonsense”. This affected 13.5% of the trials with experimental items. Repeated measures ANOVAs were computed with the two factors verb-argument-structure (three levels: missing object vs. missing subject vs. bare verb) and numerus (two levels: singular vs. plural). Two compare the two levels missing object and missing subject of verb-argument-structure, 2 × 2 repeated measures ANOVAs were computed. In case of violations of sphericity the degrees of freedom were corrected by using the Greenhouse-Geisser procedure (Huynh 1978), but the text always reports the uncorrected degrees of freedom.

For purposes of qualitative analysis four types of sensible sentence completions were classified. First, participants were using negation and n-words like niemand (nobody) to shift the achievements into states: e.g. der Bergsteiger erreichte zwei Stunden lang das Basislager nicht (for two hours the mountaineer wasn’t able to reach the base camp) or … keinen Gipfel (… no summit). Second, they were using quantified expressions like alle anvisierten Ziele (all destinations he had aimed at). Third, they used bare plural objects and/or subjects. Fourth, they changed the meaning of the verb by adding a supplementary prefix to it, for instance zwei Jahre lang fund keine Party.
The Processing of Events

6.2.2 Results

Quantitative analysis

Figure 6.2 depicts the percent of “nonsense” answers for the experimental items and the distractors. The nonsensical fillers were rejected as nonsensical in 85.5% while the sensible fillers were completed in a meaningful way in 88.7%. This indicates that participants had understood the task and provided a completion if this was possible.

Figure 6.2 Percent “nonsense” answers in Experiment 7 (+95% confidence intervals)

The experimental items were overwhelmingly rejected as nonsensical with a mean of 70.1% nonsense answers. There were, however, differences among the conditions. First of all, participants provided more sensible completions when they had to choose an object (63.8% “nonsense”) than when the subject was missing (76.3% “nonsense”). There were, however, differences among the conditions. Both show a similar pattern across the other completion types.

Rather surprisingly, there are no big differences between the singular and the plural conditions. Both show a similar pattern across the other completion types.

There were very little bare plural completions. Instead, subjects used negated sentences, quantified expressions or supplementary verb prefixes to yield the required Aktionsart of the adverbial.

Table 6.1 Distribution of different completion types in Experiment 7

<table>
<thead>
<tr>
<th></th>
<th>negation</th>
<th>quant.</th>
<th>bare plur.</th>
<th>verb prefix</th>
<th>other</th>
</tr>
</thead>
<tbody>
<tr>
<td>singular</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>miss. obj</td>
<td>31 (29.5)</td>
<td>27 (25.7)</td>
<td>14 (13.3)</td>
<td>30 (28.6)</td>
<td>3 (2.9)</td>
</tr>
<tr>
<td>miss. subj</td>
<td>64 (83.1)</td>
<td>9 (11.7)</td>
<td>0 (0)</td>
<td>4 (5.2)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>bare verb</td>
<td>26 (34.7)</td>
<td>14 (18.7)</td>
<td>2 (2.7)</td>
<td>29 (38.7)</td>
<td>4 (5.3)</td>
</tr>
<tr>
<td>plural</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>miss. obj</td>
<td>33 (28.9)</td>
<td>32 (28.1)</td>
<td>20 (17.5)</td>
<td>28 (24.6)</td>
<td>1 (0.9)</td>
</tr>
<tr>
<td>miss. subj</td>
<td>24 (35.8)</td>
<td>24 (35.8)</td>
<td>10 (14.9)</td>
<td>6 (9.0)</td>
<td>3 (4.5)</td>
</tr>
<tr>
<td>bare verb</td>
<td>16 (15.5)</td>
<td>35 (34.0)</td>
<td>19 (18.4)</td>
<td>25 (24.3)</td>
<td>8 (7.8)</td>
</tr>
<tr>
<td>total</td>
<td>194 (35.8)</td>
<td>141 (26.1)</td>
<td>65 (12.0)</td>
<td>122 (22.6)</td>
<td>19 (3.5)</td>
</tr>
</tbody>
</table>

note: absolute values (plus values in percent, computed row by row such that each row adds up to 100%), bare plur. = bare plural argument, quant. = quantificational argument, miss. obj. = missing object and miss. subj. = missing subject.

- Adding a prefix was only an option if the object was missing. This is due to the object imposing semantic restrictions on the verb. For example, adding the prefix statt to finden (find → take place) is only possible if the object denotes an event. If it is a physical object like den Schlüssel (the keys), the prefix isn’t applicable.
- The singular missing subject condition received for the most part completions with negation. How can this be explained? Adding a verb prefix was not possible

statt (finden = find, but stattfinden = take place). 96.5% of the sensible continuations could be assigned to one of these four categories.

Second, the interaction between verb-argument-structure and numerus was significant \( F(2, 118) = 5.46; \) MSE = .18; \( p < .05 \); \( F(2, 58) = 4.57; \) MSE = .09; \( p < .05 \). The interaction was due to the bare verb conditions receiving more completions when the verb was plural than when it was singular \( t_1(59) = 4.28; \) \( p < .01 \); \( t_1(29) = 2.95; \) \( p < .01 \), but the missing object and missing subject condition not showing a numerus effect. The main effect of numerus was not reliable \( F(1, 59) = 3.64; \) \( p = .06 \); \( F(1, 29) = 2.02; \) \( p = .17 \).

Qualitative analysis

Table 6.1 shows the distribution of the four types of completions across conditions. I will just focus on some interesting aspects about the distributional pattern without going through the table systematically.

- Adding a prefix was only an option if the object was missing. This is due to the object imposing semantic restrictions on the verb. For example, adding the prefix statt to finden (find → take place) is only possible if the object denotes an event. If it is a physical object like den Schlüssel (the keys), the prefix isn’t applicable.
- The singular missing subject condition received for the most part completions with negation. How can this be explained? Adding a verb prefix was not possible
because of the semantic restrictions imposed by the object which went with the verb. A bare plural was also not possible because the verb was singular. The same is true for quantified expressions like alle or viele (all, many) which also require a plural verb. These were the most common quantifiers that were used in the completions. It is thus quite plausible that negation was the last resort to achieve a meaningful completion.

6.2.3 Discussion

The present experiment investigated whether readers can predict forthcoming arguments that shift the Aktionsart of a yet incomplete verb-argument-structure in accordance with the input requirements of an aspectually mismatching adverbial. The findings clearly indicate that this is not the case. The initial part of sentences containing an achievement which is modified by a for-adverbial were overwhelmingly judged as nonsensical. This provides evidence for context-dependent parsing which assumes that readers just predict lexical material on an associative basis without deep aspectual analysis.

The predictive capabilities depend on the parts of the verb-argument-structure that have been encountered. Participants were able to come up with a sensible continuation more easily when the object than when the subject was missing. Although both, the internal and the external argument matter with respect to Aktionsart, the internal argument seems to be more accessible than the external argument.

Interestingly, the numerus information of the verb did not have a big influence on the ability to predict material that is yet to come. Although in the plural version of the missing-subject condition readers had to predict a plural subject phrase, they produced very little (only 10 out of 300) bare plural completions. Thus, even with supportive morphological information there was no evidence of charitable parsing.

There was no time pressure to provide a completion. During ordinary reading, however, the processor is forced to decide much faster on the interpretation of the incoming material. Thus, if there is no charitable parsing in an offline task like the one employed here it is even less likely that the processor will engage in charitable parsing during ordinary comprehension. Thus, assuming incremental aspectual parsing along the lines of the IAIH, readers can be expected to have problems with mismatching aspectual information as soon as they encounter it.

6.3 The bare verb as processing domain: Experiment 4b

Does the processor determine Aktionsart on the sole basis of the verb? Experiment 4b investigated this hypothesis by measuring reading times at adverbials that either matched the Aktionsart of achievement verbs, called for additive coercion or were aspectually mismatching. Processing was studied at a point where the direct object of the unambiguously transitive verb was still missing. The experiment was part of Experiment 4 which also tested adverbial modification of complete verb-argument structures. This experiment (Experiment 4a) provided evidence of processing difficulty both in additive coercion and in aspectual mismatch. Crucially, in Experiment 4a difficulty was found right at the critical adverbial.

Experiment 4b tested adverbial modification in the configuration of a subject that is followed by a transitive achievement-verb. Under standard assumptions about the way compositional interpretation of the sentence works the subject cannot be combined with a transitive verb before the direct object is present (Heim & Kratzer 1998). Consider the first words of a simple sentence in (12a) with the simplistic semantic representation in (12b).

(12)  a. John reaches …

b. Functional application of the subject node and the verb node is not possible before the VP node is semantically determined. But this depends on the object. As a result, composition has to wait until the object is present. In (13), only after the object has been encountered the interpretations of the VP and the sentence can be derived.

(13)  a. John reaches the top of K2.

b. Are these hierarchical dependencies also reflected in the processing of sentential semantic phenomena like aspectual coercion and aspectual mismatch? If they are,

3. However, Heim & Kratzer (1998) actually argue for top down interpretation which is somewhat different from the bottom approach chosen here. What is crucial here is that top down interpretation also requires a complete sentence to compute a meaning for it.
asymmetries between the processing of extraposed subjects and objects are expected. In the present experiment, I tested the aspectual status of transitive achievement verbs which haven’t received an object yet. According to semantic theory, this case should be processed exactly like the bare verb without any arguments. If we find immediate processing difficulty in the present study, an interesting next step will be to extend the paradigm to the processing of adverbial-verb-constructions. On the other hand, if we don’t find difficulty, it might be the case that Aktionsart only comes into play at the VP level. This will be investigated in Experiment 8 which will be the topic of the next section.

6.3.1 Method

Experiment 4b was part of the same experiment as Experiment 4a. The experiment used a 3 × 2 factorial design with the factors \textit{aspect} (3 levels: \textit{control} vs. \textit{additive coercion} vs. \textit{mismatch}) and \textit{word order} (two levels: \textit{svoa} (= Exp. 4a) vs. \textit{svao} (= Exp. 4b)). Here, we will only focus on Exp. 4b. (14) is a sample item. (14a) is the aspectual control condition, (14b) involves additive coercion and (14c) is an instance of aspectual mismatch. The full list of items is provided in Appendix G.

\begin{enumerate}
\item[(14)] a. Der Förster entdeckte vor zehn Min. im Wald die Falle für Bären.
   The ranger spotted ten min. ago in the forest the trap for bears.
\item b. Der Förster entdeckte in zehn Min. im Wald die Falle für Bären.
   The ranger spotted in ten min. in the forest the trap for bears.
\item c. Der Förster entdeckte zehn Min. lang im Wald die Falle für Bären.
   The ranger spotted for ten min. in the forest the trap for bears.
\end{enumerate}

The items were constructed with two spillover regions. The adverbial was followed by a prepositional phrase which was divided into two regions, the preposition and the rest of the PP. Only then followed the direct object. An effect of \textit{aspect} at the direct object region is thus very unlikely to be a spillover effect from the adverbial region. Following the object, another PP was included as second spillover region. Like the first PP, it was divided into two segments. It was always attached to the object to make the noun phrase heavier and thus more natural in extraposed position (this is a case of “heavy NP shift” see e.g. Culicover & Jackendoff (2005)). To simplify matters, statistical analyses used reading times that were aggregated over the two PP segments.

6.3.2 Results

Sensicality ratings were similar to those of Experiment 4a. Figure 6.3 depicts the percent of “yes, makes sense” judgments of both experiments.

![Figure 6.3 Sensicality judgments of Experiment 4a in black and of Experiment 4b in grey (error bars indicate 95% confidence intervals)](image)

Although the patterns are similar, mismatch detection was better in Experiment 4a than in 4b. Also, additive coercion was judged acceptable less often in Experiment 4a than in 4b. However, a direct comparison between the judgment results of Experiment 4a and 4b is difficult, because the experimental items in the two conditions differed in length and furthermore the items of Experiment 4b involved an additional PP which the items of Experiment 4a did not.

In Experiment 4b, control was judged sensible in 90.7%, additive coercion in 63.3% and aspectual mismatch in 29.9%. ANOVAs comparing the judgments of Experiment 4b revealed a significant main effect of \textit{aspect} ($F_{1,258} = 61.35; \text{MSE} = 2.88; p < .01$; $F_{2,258} = 51.00; \text{MSE} = 3.06; p < .01$). Judgments were on average provided in 1523 ms in the mismatch condition, in 1448 ms in the coercion condition and in 1510 ms in the control condition. ANOVAs revealed that judgment times didn’t differ across conditions ($F_{1,2} < 1$).

Figure 6.4 shows mean reading times of sentences with extraposed objects across the three conditions. The conditions did not differ either at the adverbial region ($F_{1,2} < 0.5$) nor at the following segments (all $F_{1,2} < 1.5$; all $F_{2} < 1$).
The Processing of Events

clearly had accomplished an aspectual interpretation. Judgments were relatively fast of the sentence, when participants were asked to provide a sensicality judgment, they determined and aspectual interpretation seemed to be delayed. However, at the end immediately interpret the initial part of the achievements, Aktionsart was not immediately (1987); Crocker (1996); Hagoort (2003a)). Although readers could in principle imme
down reading pace significantly. Similarly, in the experiments in Chapter 4 imme-
stream the sentence when readers encountered the extraposed object? A possible explanation for the lack of effect may be that the materials contained adjuncts – the first spillover region – that intervened between the adverbial and the direct object. Although they were kept constant across conditions, the intervening material may have slowed down processing of the following material in general. In turn, potential aspectual effects may have been obscured. In fact, there is psycholinguistic evidence for difficulty caused by intervening adjuncts. For instance, in a study on rightward dislocated arguments by Staub, et al. (2006) the authors found longer reading times at the adjunct PP from the stands for difficulty caused by intervening adjuncts. For instance, in a study on rightward dislocated arguments by Staub, et al. (2006) the authors found longer reading times at the adjunct PP from the stands for difficulty caused by intervening adjuncts. For instance, in a study on rightward dislocated arguments by Staub, et al. (2006) the authors found longer reading times at the adjunct PP from the stands for difficulty caused by intervening adjuncts. For instance, in a study on rightward dislocated arguments by Staub, et al. (2006) the authors found longer reading times at the adjunct PP from the stands for difficulty caused by intervening adjuncts. For instance, in a study on rightward dislocated arguments by Staub, et al. (2006) the authors found longer reading times at the adjunct PP from the stands for difficulty caused by intervening adjuncts. For instance, in a study on rightward dislocated arguments by Staub, et al. (2006) the authors found longer reading times at the adjunct PP from the stands for difficulty caused by intervening adjuncts. For instance, in a study on rightward dislocated arguments by Staub, et al. (2006) the authors found longer reading times at the adjunct PP from the stands for difficulty caused by intervening adjuncts. For instance, in a study on rightward dislocated arguments by Staub, et al. (2006) the authors found longer reading times at the adjunct PP from the stands for difficulty caused by intervening adjuncts. For instance, in a study on rightward dislocated arguments by Staub, et al. (2006) the authors found longer reading times at the adjunct PP from the stands for difficulty caused by intervening adjuncts. For instance, in a study on rightward dislocated arguments by Staub, et al. (2006) the authors found longer reading times at the adjunct PP from the stands for difficulty caused by intervening adjuncts. For instance, in a study on rightward dislocated arguments by Staub, et al. (2006) the authors found longer reading times at the adjunct PP from the stands for difficulty caused by intervening adjuncts. For instance, in a study on rightward dislocated arguments by Staub, et al. (2006) the authors found longer reading times at the adjunct PP from the stands for difficulty caused by intervening adjuncts. For instance, in a study on rightward dislocated arguments by Staub, et al. (2006) the authors found longer reading times at the adjunct PP from the stands for difficulty caused by intervening adjuncts. For instance, in a study on rightward dislocated arguments by Staub, et al. (2006) the authors found longer reading times at the adjunct PP from the stands for difficulty caused by intervening adjuncts. For instance, in a study on rightward dislocated arguments by Staub, et al. (2006) the authors found longer reading times at the adjunct PP from the stands for difficulty caused by intervening adjuncts. For instance, in a study on rightward dislocated arguments by Staub, et al. (2006) the authors found longer reading times at the adjunct PP from the stands for difficulty caused by intervening adjuncts. For instance, in a study on rightward dislocated arguments by Staub, et al. (2006) the authors found longer reading times at the adjunct PP from the stands for difficulty caused by intervening adjuncts. For instance, in a study on rightward dislocated arguments by Staub, et al. (2006) the authors found longer reading times at the adjunct PP from the stands for difficulty caused by intervening adjuncts. For instance, in a study on rightward dislocated arguments by Staub, et al. (2006) the authors found longer reading times at the adjunct PP from the stands for difficulty caused by intervening adjuncts. For instance, in a study on rightward dislocated arguments by Staub, et al. (2006) the authors found longer reading times at the adjunct PP from the stands for difficulty caused by intervening adjuncts. For instance, in a study on rightward dislocated arguments by Staub, et al. (2006) the authors found longer reading times at the adjunct PP from the stands for difficulty caused by intervening adjuncts. For instance, in a study on rightward dislocated arguments by Staub, et al. (2006) the authors found longer reading times at the adjunct PP from the stands for difficulty caused by intervening adjuncts. For instance, in a study on rightward dislocated arguments by Staub, et al. (2006) the authors found longer reading times at the adjunct PP from the stands for difficulty caused by intervening adjuncts. For instance, in a study on rightward dislocated arguments by Staub, et al. (2006) the authors found longer reading times at the adjunct PP from the stands for difficulty caused by intervening adjuncts. For instance, in a study on rightward dislocated arguments by Staub, et al. (2006) the authors found longer reading times at the adjunct PP from the stands for difficulty caused by intervening adjuncts. For instance, in a study on rightward dislocated arguments by Staub, et al. (2006) the authors found longer reading times at the adjunct PP from the stands for difficulty caused by intervening adjuncts. For instance, in a study on rightward dislocated arguments by Staub, et al. (2006) the authors found longer reading times at the adjunct PP from the stands for difficulty caused by intervening adjuncts. For instance, in a study on rightward dislocated arguments by Staub, et al. (2006) the authors found longer reading times at the adjunct PP from the stands for difficulty caused by intervening adjuncts. For instance, in a study on rightward dislocated arguments by Staub, et al. (2006) the authors found longer reading times at the adjunct PP from the stands for difficulty caused by intervening adjuncts. For instance, in a study on rightward dislocated arguments by Staub, et al. (2006) the authors found longer reading times at the adjunct PP from the stands for difficulty caused by intervene
In the next section I will present a reading time study investigating extraposed subjects which did not use any intervening adjuncts between. We will see clear indication of delayed effects there.

6.4 The VP as processing domain: Experiment 8

Does the verb plus its internal argument form a natural unit with respect to Aktionsart? Intuitive judgments reveal that the VP already encodes a minimal situation. For instance, we can talk about situations like \[\text{VP to build a house}\] while we leave it open who is actually building it. Similarly, passive sentences like \[\text{the plane has crashed}\] express complete situations.

Infinitival or gerundive VPs offer linguistic arguments for the claim that VPs express situation types. Consider (16).

(16) a. John \[\text{started to run/running}.\]
b. #John \[\text{started to win/winning the race}.\]

In (16a) to run and running respectively express an activity with which a start event can be easily combined. In contrast, in (16b) the achievement \[\text{win the race}\] has no preparatory process and the start event cannot be connected to a starting process rendering the sentence semantically anomalous. Thus, the contrast between (16a) and (16b) suggests that even subjectless VPs carry aspectual information.

But cases like these raise the problem of control. This problem concerns how to determine the understood subject of infinitival or gerundive VPs that lack an overt local subject (e.g. Chomsky (1981), for a different view see Culicover & Jackendoff (2005, Chapter 12)). The examples in (17) make clear, though, that actually no local subject is needed.

(17) a. Es wurde begonnen den Schlüssel zu suchen.
It was begun the keys to search
Somebody began to search the keys
b. #Es wurde begonnen den Schlüssel zu finden.
It was begun the keys to find
Somebody began to find the keys

Begin states that there was a start event of some process. In (17a) \text{search the keys} is of the required type, but the achievement \text{find the keys} in (17b) is not. Crucially, the constructions in (17) are in passive voice and the expletive \text{it} only serves as a dummy subject which lacks any semantic content.

Given these linguistic facts, it is quite plausible to assume that the processor determines Aktionsart at the level of the VP. This is stated in the Complete Verb Phrase Hypothesis in (18) which is repeated from Chapter 3.

(18) Complete Verb Phrase Hypothesis (CVPH)

Only a verb with all its internal arguments is specified for Aktionsart.

The CVPH stands in sharp opposition to other linguistic facts. At the beginning of this chapter it was already demonstrated that the right choice of subject bears an important influence on the Aktionsart of the whole sentence (consider \text{visitors arrived all night}). At first sight, these linguistic facts are providing conflicting evidence. On the one hand, the VP seems to be sufficient to determine Aktionsart, but on the other hand, complete verb-argument structures have to be considered.

6.4.1 Method

Materials and procedure

The present reading time experiment tested the CVPH using slightly modified materials of Experiment 4 with extraposed subject phrases. (19) is a sample item, slashes indicate segmentation. See Appendix J for the whole set of items.

(19) a. Den Haarriss/ am Wasserrohr/ bemerkte/ vor dreißig Minuten/ …
The hairline crack at the water-pipe noticed thirty min. ago …
Thirty min. ago, […] noticed the hairline crack at the water-pipe.
b. Den Haarriss/ am Wasserrohr/ bemerkte/ in dreißig Minuten/ …
The hairline crack at the water-pipe noticed in thirty min. …
In thirty min., […] noticed the hairline crack at the water-pipe.
c. Den Haarriss/ am Wasserrohr/ bemerkte/ dreißig Minuten lang/ …
The hairline crack at the water-pipe noticed for thirty min. …
For thirty min., […] noticed the hairline crack at the water-pipe.
ein aufmerksamer/ Klempner.
an attentive plumber

(19a) is aspectual control, (19b) involves additive coercion and (19c) contains an aspectual mismatch. The case disambiguated object always appeared in the sentence initial position. To license the object in that position it was always definite and maximally specific. Further, to make the object-initial word order expected, all sentences in the experiment, items and fillers, had an object before subject word order.

The numerus of the verb may provide some information about the forthcoming subject. A bare plural subject, for instance, is impossible following a singular verb. For this reason, besides \text{aspect, numerus} was manipulated in a factorial design resulting in a total of six conditions. Each item in each aspectual condition was constructed in
two versions, with a singular subject (like an attentive plumber) and with a plural subject (like a few attentive plumbers).

The 75 fillers from Experiment 4 were transformed into object-initial sentences. Items and fillers were assigned to six lists according to a latin square design. The experimental procedure was identical to Experiment 4.

Participants
42 native German speakers (31 female; mean age 23.0 years, min. 19, max. 31) recruited at Tübingen University campus took part in the experiment for 5 Euro cash. Participants were randomly assigned to lists (7 subjects per list).

6.4.2 Results
Judgments
Sensicality judgments are depicted in Figure 6.5 showing aggregated judgments over the singular (sing.) and plural (pl.) conditions.

![Figure 6.5 Mean of "yes, makes sense" judgments (+95% CIs) in Experiment 8](image)

While control was accepted in 89.7% (sing.: 92.9% vs. pl.: 86.6%), mismatch was only accepted in 31.3% (sing.: 26.8% vs. pl.: 35.7%). Coercion was intermediate with 63.0% "yes" responses (sing.: 58.7% vs. pl.: 67.6%). The sentences involving aspectual coercion were judged as sensible in the majority of cases, as was confirmed by a t-test testing whether coercion were judged as sensible in the majority of cases, as was confirmed by a t-test. While the main effect of numerus was not reliable ($F_{1, 41} = 3.25; MSE = 0.10; p = .08$; $F_{1, 29} = 2.07; MSE = 0.07; p = .17$), the interaction between numerus and aspect was significant ($F_{1, 2, 82} = 5.32; MSE = 0.17; p < .05$; $F_{2, 2, 58} = 4.18; MSE = 0.12; p < .05$). The interaction is due to the fact that the differences between the aspectual conditions are bigger in the singular than in the plural conditions.

Table 6.2 shows the judgment times for "no" responses in the mismatch conditions and "yes" responses in the coercion and control conditions.

Table 6.2 Judgment times in Experiment 7

<table>
<thead>
<tr>
<th>Condition</th>
<th>Judgment time (+ std.)</th>
<th>N (out of 210)</th>
</tr>
</thead>
<tbody>
<tr>
<td>singular</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&quot;no&quot; to mismatch</td>
<td>1254 (664)</td>
<td>154</td>
</tr>
<tr>
<td>&quot;yes&quot; to coercion</td>
<td>1549 (759)</td>
<td>121</td>
</tr>
<tr>
<td>&quot;yes&quot; to control</td>
<td>1341 (633)</td>
<td>195</td>
</tr>
<tr>
<td>plural</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&quot;no&quot; to mismatch</td>
<td>1294 (655)</td>
<td>134</td>
</tr>
<tr>
<td>&quot;yes&quot; to coercion</td>
<td>1495 (666)</td>
<td>142</td>
</tr>
<tr>
<td>&quot;yes&quot; to control</td>
<td>1331 (672)</td>
<td>181</td>
</tr>
</tbody>
</table>

Note: Judgment times of expected answers in ms.

In both numerus conditions, judgments took longer for sentences involving aspectual coercion than for controls or sentences involving an aspectual mismatch. In ANOVAs which analyzed judgment times of expected answers (= "no" with respect to mismatch and "yes" with respect to coercion and control), this difference was reflected by a main effect of aspect that was significant by subjects ($F_{1, 2, 82} = 5.14; MSE = 689746.93; p < .05$; $F_{2, 2, 58} = 2.84; MSE = 530684.61; p = .08$). Neither the main effect of numerus ($F_{1, 2, 82} < 1$) nor the interaction between aspect and numerus was reliable ($F_{1, 2, 58} < 1$). A paired t-test comparing judgment times for coercion versus control (pooled over numerus conditions) revealed that the difference between them was reliable ($t_{1, 41} = 2.40; p < .05$; $t_{2, 29} = 3.10; p < .01$).

Reading times
The reading times for the three aspectual conditions are depicted in Figure 6.6. They were longer in the aspectual mismatch and the coercion condition compared to control. Since the pattern was the same in the singular and the plural conditions, data were aggregated over the corresponding singular and plural conditions.

---

5. Three (mismatch vs. coercion vs. control) by two (singular vs. plural) ANOVAs for the segments adverbial, first part of the subject and second part of the subject revealed no significant main effects and interactions involving numerus. Thus, in the following, only effects of aspect will be reported.
The difference in reading times only showed up at the head noun of the extraposed subject phrase (mismatch: 88.3 ms/char vs. coercion: 88.1 ms/char vs. control: 77.3 ms/char).

At the adverbial region, the aspectual conditions didn’t differ in reading time. ANOVAs didn’t reveal a significant main effect of aspect ($F_1(2, 82) = 2.64; \text{MSE} = 253.02; p = .09; F_1(2, 58) = 1.01; \text{MSE} = 180.68; p = .36$).

At the following first part of the subject phrase, there were also no differences in reading times between conditions. Numerically, control was even slowest. ANOVAs didn’t reveal a significant main effect of aspect ($F_1(2, 82) = 2.50; \text{MSE} = 467.97; p = .09; F_1(2, 58) = 1.70; \text{MSE} = 376.71; p = .20$).

When readers encountered the head noun of the subject phrase, reading times were slower in the mismatch and the coercion conditions than in the singular and plural controls. In ANOVAs, this difference was reflected by a significant main effect of aspect ($F_1(2, 82) = 7.32; \text{MSE} = 4880.34; p < .01; F_1(2, 58) = 7.88; \text{MSE} = 3801.87; p < .01$).

### 6.4.3 Discussion

In the present reading time study participants were reading sentences with extraposed subject phrases involving aspectual mismatch and aspectual coercion. Sensicality judgments revealed that readers paid attention to the aspectual properties of the sentences. Aspectual mismatch was rated as nonsensical, aspectual controls were judged to be semantically well formed and aspectual coercion sentences were accepted in the majority of cases.

Reading times indicate difficulty in processing sentences that involve aspectual mismatch and coercion compared to aspectual control. Interestingly, this difficulty only emerged after readers had encountered the extraposed subject phrase, that is only at the point when the verb argument structure is complete. In contrast, at the critical adverbial and the subsequent region all three conditions were read equally fast. The results thus provide clear evidence against the CVPH. The VP did not contain enough information to allow for aspectual mismatch and coercion effects when it was combined with a mismatching or coercing adverbial. Further, since only delayed effects were found, it is not surprising that the numerus information wasn’t used to predict what kind of subject is yet to come.

In contrast to the findings of Experiment 4b, the results of the present study show delayed aspectual effects. When the verb had received all its arguments, the adverbial was semantically combined with the complete verb-argument-structure as indicated by processing difficulty in response to mismatch and coercion. This delayed effect calls for a hierarchical organization of aspectual processing, where first the eventuality of the verb-argument structure has to be computed. Only then adverbial modification is computed. The findings thus support the Late Aspectual Interpretation Hypothesis (LAIH).

Aspectual processing difficulty started at the extraposed subject phrase. But it went on way beyond the sentence and even affected the judgment times of coerced sentences. Mismatch detection, for which judgment times were roughly at the same level as control, elicited only a local effect at the subject phrase. The different time course of computing aspectual mismatch and additive coercion hints at functional differences between the two conditions. I will come back to this issue in the ERP study in the next chapter.

More generally, the findings demonstrate a fascinating interplay between the parsing of argument structure and of Aktionsart. The former seems to be prior to aspectual processing which depends on an analysis of the argument structure. This adds an interesting new parameter to the incrementality debate, namely the domain size with respect to a particular phenomenon.

The important role of argument structure becomes even more obvious if we compare the findings of Experiments 4b and 8 with Experiment 2. In Experiment 2 processing difficulty emerged at in-adverbials which followed a minimal intransitive activity sentence (e.g. Peter jogged…) which allows for additional material that can shift the Aktionsart. In contrast, Experiments 4b and 8 used achievements which were unambiguously transitive. It thus seems as if aspectual processing makes a clear difference between constructions which allow both for intransitive and transitive uses and unambiguous transitive constructions.
6.5 Summary and conclusions

In the present chapter I investigated the processing domain of Aktionsart. I laid out two hypotheses, *incremental aspectual interpretation* (IAIH) versus *late aspectual interpretation* (LAIH). The first hypothesis is inspired by psycholinguistic work on syntactic processing which shows that the syntactic representation is constructed on a word-by-word basis. The LAIH takes into account semantic work on Aktionsart like Dowty (1979), Verkuyl (1993) and Krifka (1998) which demonstrates that Aktionsart can only be determined at the sentential level.

In two reading time studies (Experiments 4 and 8) I used adverbial modification of yet incomplete verb-argument structures to investigate whether aspectual mismatch and additive coercion slow down reading of the adverbial when arguments are still missing. The results of both experiments provide evidence for the LAIH: the adverbial only showed semantic effects after the verb had received all its’ arguments. These findings are particularly striking since a completion study (Experiment 7) showed that readers were judging the same fragments as semantically ill-formed which they were reading at normal pace during ordinary comprehension.

The findings are inconsistent with the IAIH. Aktionsart seems to be determined at the sentence level at the earliest. Does this mean that semantic phenomena like Aktionsart aren’t processed incrementally? Reflecting upon the notion of *incrementality*, two senses can be distinguished. First, *incrementality* sometimes means *immediacy* which reflects whether some kind of information is taken into account immediately, that is during first interpretation. Second, *incremental interpretation* sometimes is used to refer to processing that proceeds *word-by-word*. In principle, these two aspects are independent from each other and have to be kept apart.

Processing can proceed in a *word-by-word* fashion although it may not be *immediate*. An actual example is provided by modular theories about *sentence processing* (e.g. Fodor (1983); Frazier (1987); Friederici (2002)). According to these theories, the immediate interpretation is solely built upon syntactic information while other sources of information (e.g. semantic) are only considered at later stages. These later stages may well be performed while the perceiver is still at the same word.

On the contrary, processing can be *immediate*, but not computed on a *word-by-word* basis. For many semantic phenomena this seems to be a quite plausible option. Consider for instance a subsective interpretation of an adjective like in *beautiful dancer* (for intersective vs. subsective interpretations see e.g. Kamp & Partee (1995)). What should the processor do when encountering the adjective? Is *beautiful* first interpreted, say intersectively, and then reanalyzed when *dancer* is encountered? Interestingly, in an unpublished paper by Frisson, et al. (ms.) found no differences in first pass reading times between *heavy smoker* and *heavy teenager* at the noun, although they used biasing context which should enforce the subsective or the intersective reading, respectively.

Thinking a bit about other types of semantic phenomena, *word-by-word* processing seems often to be not very likely. How, for instance, are negative polarity items (e.g. *any*) processed if the required negative context hasn’t been established yet? What about idioms like *kick the bucket*? Are these first interpreted literally causing a local breakdown of interpretation? These phenomena suggest that semantic processing must rely upon bigger chunks of language than is required for syntactic processing. If we assume a larger processing domain of aspectual interpretation the question remains whether aspectual processing is immediate in the sense above. This will be the issue of the next chapter.
CHAPTER 7

The processing of temporality in the brain

In this chapter the processing of temporality will be studied using neurophysiological measures. I will report an EEG study which investigated the brain responses of local and global aspectual mismatch. These measures will be used to decide between different theoretical accounts of aspectual coercion. Furthermore, different aspects of temporality will be compared: Aktionsart versus tense.

Section 7.1 serves as an introduction to EEG research in psycholinguistics. Section 7.2 provides a brief overview over brain studies on temporality in language. In Section 7.3 I will report an ERP study on the processing of Aktionsart (Experiment 9). Section 7.4 presents a summary and conclusions.

7.1 EEG research in psycholinguistics

In psycholinguistic studies using electroencephalographic (EEG) data the ongoing EEG is recorded while a participant reads or hears a piece of language. Typically, the researcher is only interested in those parts of the EEG that reflect the processing of the linguistic stimulus. The activity evoked by the stimulus leads to changes in the brain potentials that are time-locked to its onset and is measured using event related potentials (ERPs). The ERP methodology has excellent temporal resolution and allows to discriminate between functionally different processes during sentence comprehension.

I will provide a short introduction to the measurement of ERPs in turn.

7.1.1 Measuring event-related brain potentials

ERPs are small changes in the electrical activity of the brain which occur in response to cognitive events and are obtained by using electrodes placed on the scalp. ERPs are a non-invasive measurement technique. A simple experiment studying the processing of frequent versus infrequent stimuli is illustrated in Figure 7.1.

While a participant reads or hears a sentence, the ongoing EEG is recorded. The potential changes evoked by the stimuli are much smaller than the background activity
The Processing of Events

of the brain which is random noise relative to the evoked activity. By averaging portions of the EEG time-locked to the onset of repeated stimuli of the same type the ERPs can be extracted from the background noise.

The basic assumption underlying ERP research is that different types of experimental manipulations elicit different patterns (so-called components) of brain activity. An ERP-waveform is thereby taken as a composite of different components that are

superimposed on each other. ERP components can be characterized along the following four dimensions: polarity (positive vs. negative), latency (the time involving onset, peak and duration at which the effect is visible), amplitude (the strength of the effect) and topography (the distribution over the scalp). Typically, ERP components are named by their polarity and peak latency (e.g. P100, P300, N400 and P600). Roughly speaking, the early components (like the P100) reflect exogenous processes and are dependent on the physical properties of the stimulus, such as luminance. Endogenous components, in contrast, depend entirely on the task performed by the subject and are not directly influenced by the physical properties of the eliciting stimulus (see e.g. Rugg & Coles (1995); Luck (2005)). Like reaction times ERPs only provide relative
The Processing of Events

simple subject-predicate statements using sentence verification tasks. For instance, the N400 seems to be unaffected by sentential negation in studies examining created by a preceding sequence of words or line drawings (Ganis, et al. 1996; Holcomb & example a line drawing will elicit an N400 if it is inconsistent with the semantic context word) or in lexical frequency (like isolated words give rise to an N400 if they differ in lexical class (content vs. function word) or in lexical frequency (van Petten & Kutas 1990), repetition (van Petten, et al. 1991) and discourse context (van Berkum, et al. 1999b). Two EEG studies on complement coercion by Choma (2006) and by Kuperberg et al. (2009) found an N400 effect of comparable size for both complement coercion and semantic mismatch compared to a control construction. Since in the semantic litera
ture both complement and aspectual coercion have been treated in a uniform manner (Dölling 2003a) it is an especially interesting question if aspectual coercion leads to an N400, too.

Though the integration of a content word generally gives rise to the N400, it becomes more pronounced with a higher degree of unexpectedness of the critical word (Kutas & Hillyard 1984). In addition, different types of lexical-semantic manipulations modulate the N400, for instance lexical frequency (van Petten & Kutas 1990), repetition (van Petten, et al. 1991) and discourse context (van Berkum, et al. 1999b). Two EEG studies on complement coercion by Choma (2006) and by Kuperberg et al. (2009) found an N400 effect of comparable size for both complement coercion and semantic mismatch compared to a control construction. Since in the semantic litera
ture both complement and aspectual coercion have been treated in a uniform manner (Dölling 2003a) it is an especially interesting question if aspectual coercion leads to an N400, too.

The N400 is also observed for words without sentence context. For instance in a word list the N400 gets more pronounced the less expected a new lexical item is. Even isolated words give rise to an N400 if they differ in lexical class (content vs. function word) or in lexical frequency (like monocle vs. milk).

Nonlinguistic stimuli can also elicit an N400 as long as they are meaningful. For example a line drawing will elicit an N400 if it is inconsistent with the semantic context created by a preceding sequence of words or line drawings (Ganis, et al. 1996; Holcomb & McPherson 1994).

Is the N400 also modulated by semantic processing at the phrasal level? Interest-
ingly, the N400 seems to be unaffected by sentential negation in studies examining simple subject-predicate statements using sentence verification tasks. For instance,

The N400

The best studied language-related component is the N400, first reported by Kutas & Hillyard (1980). The N400 is a negative-going wave peaking around 400 ms post stimulus and is usually largest over central and parietal electrode sites, with a slightly larger amplitude over the right than over the left hemisphere. The N400 is typically observed in response to violations of lexical semantic expectations. In the field of sentence processing, the N400 is interpreted as reflecting difficulties with the integration of a new word in the current semantic representation of the sentence. A typical example is provided in (1) where an N400 effect can be observed when the verb cry is encountered in (1a) compared to eat in (1b).

(1) a. The pizza was too hot to cry.
   b. The pizza was too hot to eat.

Though the integration of a content word generally gives rise to the N400, it becomes more pronounced with a higher degree of unexpectedness of the critical word (Kutas & Hillyard 1984). In addition, different types of lexical-semantic manipulations modulate the N400, for instance lexical frequency (van Petten & Kutas 1990), repetition (van Petten, et al. 1991) and discourse context (van Berkum, et al. 1999b). Two EEG studies on complement coercion by Choma (2006) and by Kuperberg et al. (2009) found an N400 effect of comparable size for both complement coercion and semantic mismatch compared to a control construction. Since in the semantic litera
ture both complement and aspectual coercion have been treated in a uniform manner (Dölling 2003a) it is an especially interesting question if aspectual coercion leads to an N400, too.

The N400 is also observed for words without sentence context. For instance in a word list the N400 gets more pronounced the less expected a new lexical item is. Even isolated words give rise to an N400 if they differ in lexical class (content vs. function word) or in lexical frequency (like monocle vs. milk).

Nonlinguistic stimuli can also elicit an N400 as long as they are meaningful. For example a line drawing will elicit an N400 if it is inconsistent with the semantic context created by a preceding sequence of words or line drawings (Ganis, et al. 1996; Holcomb & McPherson 1994).

Is the N400 also modulated by semantic processing at the phrasal level? Interest-
ingly, the N400 seems to be unaffected by sentential negation in studies examining simple subject-predicate statements using sentence verification tasks. For instance,

The P600

A number of ERP components have been proposed for syntactic processing. The most studied is the P600, a late positive shift with a maximum over parietal and central midline electrode sites.1 Like for the N400 there is no undisputed interpretation of the P600.

The P600 was first observed in a study by Osterhout & Holcomb (1992) on the processing of garden-path sentences like in (3a) compared to an unambiguous control in (3b).

(3) a. The broker persuaded to sell the stock was sent to jail.
   b. The broker that was persuaded to sell the stock was sent to jail.

Since in garden-path sentences the preferred structure has to be abandoned in favor of an initially dispreferred interpretation, the P600 has been related to a late controlled process of syntactic repair in a multistage parsing model (Friederici & Mecklinger 1996). A second interpretation takes the P600 to be an index for the inability of the parser to assign the preferred interpretation to the input (Hagoort, et al. 1993). These two early exegeses differ in terms of the stage at which the P600 originates, but both take it to be an indication of revising an initial syntactic interpretation in case of syntactic ambiguity. Kaan et al. (2000) showed that P600 effects can also be found in unambiguous sentences. For this reason, they proposed that the P600 has to be interpreted as an index of syntactic integration difficulty in general.

1. In fact there is discussion if the P600 even may comprise a whole family of different components which can be distinguished by their temporal properties and their distribution. For instance Friederici (2002) proposes two different P600 components: one more frontal P600 reflecting syntactic integration problems and a centro-parietal P600 which indicates repair processes.
Interestingly, Kim & Osterhout (2005) found a P600 effect, but no N400 in thematically ill-formed sentences like (5a) in contrast to a control in (5c). 2

(5) a. The hearty meal was devouring the kid.  
   b. The hearty meal was devoured by the kid.  
   c. The hungry boy was devouring the cookie.

Thus, it seems that semantic violations also can produce P600 effects. The authors explain this finding by assuming that the reader implicitly carries out syntactic repair. This process is initiated by thematic properties of the sentence. Thus, readers reanalyze (5a) along the lines of (5b) to fulfill the animacy requirement of the subject.

A finding that challenges the classical interpretations of the P600 comes from the processing of so called jabberwocky sentences. In these sentences all content words are replaced by pronounceable pseudowords. If a reader encounters a syntactic violation in a jabberwocky sentence little or no P600 effect emerges, although in syntactically parallel sentences with preserved meaning a P600 effect can be observed (Monchi, et al. 1997; Huddleston, et al. 2003). The jabberwocky paradigm from Monchi et al. (1997) is illustrated in (6).

(6) a. Der Kruke plötzt den Schruck.  
   A flurk merches the minch  
   b. *Das Klenk frunen den Wech.  
   A mizzel quanch the plurr  
   c. Der Junge schlacht den Hund.  
   The boy beats the dog  
   d. *Der Mann trinken das Bier.  
   The man drink the beer

The examples (6b) and (6d) contain a congruence violation of the finite verb. The subject is singular, but the verb appears in its plural form. Although grammaticality judgments revealed that readers recognize the number mismatch in (6b), in the evoked potentials a P600 was only present in condition (6d). These findings suggest that the P600 is not triggered by purely syntactic difficulty, but also reflects semantic processing at a supralexical level. Thus, the most general interpretation is to treat the P600 as a syntactic and semantic index of integration difficulty at the phrasal level.

Moreover, some researchers doubt that the P600 is an independent component at all (Coulson, et al. 1998a). Instead they interpret it as a general purpose response to low-probability target events often associated with some form of categorization or binary decision (P3b component). This means that the P600 is a late subtype of the P300 component. The P300 can also be observed in non-linguistic tasks like the odd-ball paradigm (see Figure 7.1). Recent imagery studies, however, have shown that the P600 and the P300 have no common neural generators (Snijders, et al. 2005). It is thus justified to treat the P600 as an independent component (see also the arguments in Osterhout & Hagoort (1999)).

To conclude, the P600 has traditionally been interpreted as a neurophysiological correlate of difficulty in syntactic processing. But since there are controversial findings, I will take it as an index of interpretation difficulty in general both reflecting syntactic and semantic integration of a stimulus.

Left anterior negativities
Violations of morphosyntactic well-formedness have typically been associated with anterior negative responses. These morphosyntactically triggered responses exhibit a certain degree of variability with regard to both latency and scalp distribution. Anterior negative responses are usually left lateralized or bilaterally distributed. They are commonly referred to as a left anterior negativities (LAN) since they are typically most pronounced at left anterior sites. While the LAN often does not exhibit a clear peak its latency falls between 300 and 500 ms post-stimulus onset, but it has also been reported as early as 100 ms. Some researchers have proposed a functional distinction between an early left anterior negativity (ELAN) with a latency between 100 and 300 ms and a later latency negativity occurring between 300 and 500 ms (reserved for the LAN) although the ELAN frequently persists into the later latency window and exhibits the same distribution.

Experimental paradigms that elicit an ELAN violate expectations that the incoming word will be of a particular grammatical category. The ELAN has been taken to index an early automatic process of local phrase structure building, during which word category information is used to assign the initial syntactic structure (Friederici 2002). However, this conclusion remains controversial, since the ELAN has to date reliably been elicited under only a narrow set of conditions involving word category violations like in (7b). In this example (taken from a study by Hahne & Friederici (2002)) a noun rather than a verb is required given the context of a preposition plus an article. This leads to an ELAN in (7b) compared to the control in (7a).

(7) a. Das Brot wurde gegessen.  
    The bread has been eaten  
    b. *Das Eis wurde im gegessen.  
    The ice cream has been in the eaten

2nd proofs
Turning to the LAN component (300–500 ms) there is also debate about its functional significance. On the one hand, the LAN component has been observed in response to word category violations, but also with agreement violations in word (Hagoort & Brown 2000) and pseudoword sentences (Monchi et al. 1997) and seems to reflect morphosyntactic processing. On the other hand, left anterior negativities have also been observed as a response to sentences which contain no violations, but lead to increased working memory demands. Thus, the LAN is hypothesized to indicate a “look forward” function of displaced sentence constituents. These include for instance filler-gap constructions (e.g. King & Kutas (1995)). Alternatively, it can also indicate a “look back” if syntactic information has to be reconciled and aligned with preceding information. This is for instance the case if an anaphor is seeking an appropriate antecedent (Coulson, et al. 1998b). The morphosyntactic and the working memory LAN can be distinguished on the basis of their duration. While a morphosyntactic LAN is typically limited to the time-interval 300 to 500 ms post-onset of the stimulus, a working memory LAN is often sustained over a longer period of time. For instance, in (8a) a working memory LAN beginning 300 to 500 ms after the onset of the filler who is expected which will last until the parser eventually encounters the gap (indicated by $t_1$). Thus, in (8a) a more sustained LAN effect is expected than in (8b).

(8) a. Who, did you say that Mary kissed $t_1$?
   b. Who, $t_1$ said that Mary kissed John?

In ERP research different components can appear together in one waveform. For example filler-gap constructions typically trigger a working memory LAN which begins shortly after the onset of the filler. When the parser encounters the gap typically a P600 emerges (e.g. Phillips, et al. (2005)). Another example for combined effects are agreement violations which often elicit combined LAN-P600 patterns (e.g. Steinhauer & Ullman (2003)). Combinations of N400 and P600 have also been observed (e.g. Hagoort (2003b)).

A theoretical explanation of the different components in two alternative neurolinguistic models of sentence processing will be the topic of the next section.

7.1.2 Neurolinguistic models of sentence comprehension

I will describe two neurolinguistic models of sentence comprehension. The first is a modular multistage model (Friederici 1999; Friederici 2002) while the second is essentially nonmodular and assumes immediacy at all linguistic levels (Hagoort 2003a).

3. Note that the gap is essentially needed to determine grammatical aspects of the filler. For example in (8a) who is the direct object of the embedded clause while in (8b) who is the subject of the matrix clause.

Thus, the two models represent the two extremes in terms of modularity. It is still controversial if the parser works in a modular fashion (in the sense of Fodor (1983) and Frazier (1987)). In particular, it is under debate if the parser is modularly encapsulated and initially relies on structural information alone. I will first describe the two models and then discuss their predictions with respect to aspectual coercion.

Friederici’s neurolinguistic model of sentence comprehension

Friederici (Friederici 1999; Friederici 2002) proposes a neurocognitive model of sentence processing that involves three serially ordered stages. The model relates to the garden-path theory of sentence processing (Frazier 1978; Frazier 1987). Each processing stage can be characterized by its unique ERP correlates.

- **Phase 1.** This stage corresponds to the first pass parse and is associated with the ELAN component. The parser integrates a lexical item on the basis of morphosyntactic word category information alone. Semantic aspects are not taken into account.

- **Phase 2.** The initial parse is used for two processing routes which are processed in parallel, but independently of each other. In semantic processing the lexical information is composed with the sentence context. This is reflected by the N400 component. In the syntactic branch the parser makes use of all the morphosyntactic information that is available. This is reflected by the LAN component.

- **Phase 3.** This stage binds together the semantic and the syntactic representation of the previous stage. When the mapping fails, the syntactic structure has to be revised and repaired. This is reflected by the P600 component. Friederici differentiates between two kinds of P600s. While a frontally distributed P600 reflects syntactic complexity, a more centro-parietally distributed P600 indexes syntactic repair.

There has been evidence in favor of the model, but also against it. One prediction of the model is that uninterpretable information at an earlier stage should block processing at a later stage, since there will be nothing computed which can feed into the subsequent modules. This prediction has been tested by means of auditorily presented sentences that contained both a word category violation and a semantic violation at

4. Of course there are other models, too. In the neurolinguistic model by Bornkessel & Schlesewsky (2006), for instance, the authors also assume a modular architecture while focusing on the processing of thematic relations across languages. I will concentrate on the models by Friederici (2002) and Hagoort (2003a) because they are the loci classici of modular vs. nonmodular neurolinguistic approaches.
The N400 reflects semantic integration of a word or concept into the overall meaning representation of the preceding context. In sentence contexts it therefore indexes ease of semantic composition.

- **LAN**. The LAN has so far only been observed in response to syntactic violations. Therefore, it indexes the impossibility to attach a new item to the syntactic representation of the sentence context. This can be the case because the new item is of the wrong word category and therefore no possible attachment site exists. Alternatively, an agreement error (a feature mismatch among phrasal and lexical constituents) also makes attachment impossible.

P600. The P600 is related to the time it takes to attach new input to the preceding (sentence) context. Its amplitude is modulated by the amount of competition. It is affected by the number of alternative attachment sites in the case of syntactic ambiguity, by syntactic complexity and by semantic and pragmatic influences. Violations also result in a P600 as long as binding attempts are made.5

The model correctly predicts the findings of both studies mentioned in the previous section. If mismatching word category information precedes lexical semantic information a LAN is expected and the sentence can be rejected before the semantic information is encountered. Alternatively, if mismatching lexical semantic information precedes word category information semantic processing is expected to start even earlier than syntactic processing.

**Predictions of the two models with respect to aspectual coercion**

Before coming to the predictions it should clear that both models primarily focus on syntactic processing. It is therefore very hard (and perhaps unfair) to derive predictions concerning aspectual processing. On the other hand, if these theories are taken to be neurolinguistic models of language comprehension in general, it is interesting to see how far we get when applying them to semantics as well.

What does Friederici’s model predict concerning a sentence with aspectual coercion? Which phases should be affected? Type conflict is a form of semantic mismatch and should be detected during semantic integration in phase 2. As a result, mismatch detection is expected to lead to an N400 effect. Revision and repair are processes that occur during the third phase. Under a syntactic account of coercion an update of the syntactic structure is expected by adding an appropriate coercion operator to it. Another alternative (that cannot be addressed in Friederici’s primarily syntactic model) consists in purely semantic reanalysis. If the first alternative is correct a P600 is expected to follow the N400. Since repair that is carried out solely at a semantic level is not covered by the model no predictions can be made concerning the second alternative.

In contrast, the immediacy model makes slightly different predictions. While trying to compose the coercing stimulus with the preceding sentence a N400 effect is expected that indicates the detection of a semantic mismatch. What does this mean

5. Hagoort’s model is spelled out for syntactic parsing, so this may give rise to the wrong impression that the P600 is dedicated to syntactic processing in his account, too. This is clearly not the case. The semantic component is simply not spelled out in Hagoort (2003a).
for binding? While syntactic constraints are satisfied, there is an initial violation of semantic requirements of the coercing stimulus. Therefore besides an N400 also a P600 component is expected for aspeccial coercion compared to a control.

With these predictions in mind I conclude the introduction to EEG research in psycholinguistics. Since there is still considerable controversy about the functional interpretation of different ERP components and the architecture of the language processor in general we will have to be very careful in interpreting the functional significance of the results in the following two sections.

7.2 Relevant brain studies

Although to date there is no published EEG study on the processing of aspectual coercion, there are some studies that investigated another aspect of temporality, namely the processing of tense violations. Although these studies were mainly concerned with the morphosyntactic processing of tense morphemes they are also instructive with regard to the semantic interpretation of tense. Complement coercion has received some attention, too, and has been investigated in two studies, one MEG and one EEG experiment. In the following two sections I will summarize these studies.

7.2.1 Brain studies on temporality in language

Tense violations have received some attention in ERP studies on language processing. These studies were mainly concerned with the question whether regular and irregular forms are treated differently during online comprehension. For instance, following Pinker (1991) it has been proposed that regular forms have to be derived by the application of rules (like past-tensed verb = stem+ed) while irregular forms are stored in the lexicon. In the following I will ignore this aspect of the studies, but concentrate on the brain responses elicited by a mismatch in tense irrespective of verb type. In particular, I will focus on two questions: (1) which ERP components are commonly observed in response to a mismatch in tense and (2) when can a tense mismatch be recognized at the earliest (in terms of verb-argument-structure).

What is a tense violation?
Following Reichenbach (1947) tense can be characterized by three temporal indices: time of speech (S), event time (E) and reference time (R). Their relations on the time axis (which can be conceived of as the real numbers) determine the tense expressed by the sentence. Consider the difference between the Past Perfect (11a) and the Simple Past (11b).

(11) a. Mary had left the house.
    b. Mary left the house.

In the Past Perfect the event time is before the reference time which itself lies in the past of the time of speech (tense constraint for the Past Perfect: E < R, R < S). In the Simple Past event time and reference time coincide and lie in the past of the moment of speech (tense constraint for the Simple Past: E = R, R < S). Besides grammaticalized tense (tense morphemes and analytical constructions involving auxiliaries) there are also non-grammaticalized tensed expressions comprising adverbs and temporal adjunct clauses. These determine the reference time. And it is exactly here where the possibility of tense mismatch arises. Consider (12a) versus (12b).

(12) a. Yesterday, Mary left the house.
    b. *Tomorrow, Mary left the house.

(12a) is a semantically perfect English sentence. This is due to a satisfiable tense constraint: R < S, E = R, R < S. In contrast, (12b) expresses a contradiction and can therefore not be satisfied in any model: R > S, E = R, R < S. This provides us with a semantic notion of tense violation: a sentence contains a tense violation if it translates into a contradictory tense constraint.

ERP studies on tense violations

In a study conducted by Allen, Badecker and Osterhout (Allen, et al. 2003) syntactic tense violations were investigated using ERPs. They tested sentences like in (13).

(13) a. The man will work on the platform.
    b. The man will worked on the platform.

In the 500–900 ms time window post onset of the verb they observed a P600 effect for (13b) compared to (13a). They interpret the P600 to reflect morphosyntactic processing in response to a tense violation. Note however that there is a possible confound since work is an infinitive as it is required in this construction while worked is a finite verb. In principle, in their study the P600 might reflect a violation of finiteness rather than tense. Thus, not very much can be concluded concerning the functional correlates of tense violations or their processing domain.

An ERP study by Steinhauer and Ullman (Steinhauer & Ullman 2003) tested sentences like in (14).

(14) a. Yesterday, I sailed Diane’s boat to Boston.
    b. Yesterday, I sail Diane’s boat to Boston.

In (14) the adverb yesterday specifies a temporal frame for the occurrence of the event of the speaker sailing Diane’s boat to Boston. In (14b) the clause I sail Diane’s boat to Boston locates the event at (or in an in interval including) the moment of speech, thus outside the period denoted by yesterday. In the 400–900 ms time window post onset of the verb, tense violations elicited a LAN (400–500 ms) and a P600 (600–900 ms).
Steinhauer and Ullman interpreted these as indexes of morphosyntactic processing thereby largely ignoring the semantic side of tense violations. Unfortunately, in the constructions used in Steinhauer and Ullman’s study the tense violation occurred only after the subject and the verb had been presented. As yesterday, I sailed is already a complete English sentence the study cannot be interpreted with respect to the processing domain for tense.

Baggio (Baggio 2004) investigated tense violations from a semantic point of view. He proposes a theoretical framework building upon Reichenbach’s semantic treatment of tense. He claims that tense violations are essentially semantic and tested this in an ERP study using Dutch sentences like in (16).

(16) a. Afgelopen lente won/*wint Julian een literatuur prijs in Frankrijk.  
   Last spring won/*wins Julian a literature prize in France  
   Last spring, Julian won/*wins a literature prize in France  
   b. Julian won/*wint afgelopen lente een literatuur prijs in Frankrijk.  
   Last spring won/*wins last spring a literature prize in France  
   Last spring, Julian won/*wins a literature prize in France

Like in the preceding study Baggio chose verbs (always third person singular) that were included (won) or were not included (wint) in the time frame denoted by a temporal adverbial (afgelopen lente). Moreover, in (16a) the tense violation occurred before the verb had received its arguments. Another factor manipulated was the critical word. In (16a) the tense violation was introduced at the verb, whereas in (16b) the adverb was the critical segment. This manipulation enabled him to differentiate between morpho-syntactic and semantic tense violations. Crucially, in (16b) the critical adverbial is only inappropriate due to its semantics, but not morpho-syntactically.

Although there were numerical differences between the conditions resembling a LAN-P600 pattern on the critical segment in the violation conditions in both (16a) and (16b) the effects were not statistically reliable. This may be the case because present tensed verbs were used. Among all tensed verbs the present shows the highest variability in denotational freedom. Consider (17), an acceptable German sentence in which the historical present tense is present. If this is correct tense is processed fairly independently from Aktionsart (see Chapter 6 for the domain of aspectual processing). Such a finding would be particularly interesting because in temporal semantics there has been effort to treat different aspects of temporality in a uniform way (Steedman 2002; Hamm & van Lambalgen 2005).

7.2.2 Brain studies on complement coercion

Besides brain studies on temporality there have been some studies looking into the processing of semantic type shifts. All except for Brennan & Pylkkänen (2008a) investigated the processing of complement coercion. The first study by Pylkkänen et al. (Pylkkänen & McElree 2007) used MEG and two other studies by Choma (2006) and Kuperberg et al. (2007) used EEG data within the same experimental design. Sample materials are provided in (18).

(18) a. The journalist wrote the article after his coffee break.  
   b. The journalist began the article after his coffee break.  
   c. The journalist ate the article after his coffee break.

The difference between the three sentences in (18) is that in the control in (18a) the activity that the journalist is performing on the article is explicitly encoded in the
sentence, whereas in the coering case in (18b) it is left up to the reader to infer the missing action. (18c) contains a semantically inappropriate verb and provides a case of unresolvable semantic mismatch. In terms of semantic operations needed in these three cases (18a) is derived strictly compositionally, (18b) can only be interpreted via enriched composition (Jackendoff 1996) and (18c) has no literal interpretation at all.

In the study by Pylkkänen & McElree (2007) the authors found an M350 component which is an MEG correlate of a subcomponent of the N400 for the mismatch condition (18c) compared to the control in (18a). Interestingly, in the coercion condition Pylkkänen et al. found increased activation in the anterior frontal cortex at 350–450 ms (Anterior Midline Field, AMF in short) whereas anomaly affected activation in the left temporal lobe in a slightly earlier time window. The authors interpret these findings as suggesting that the increased activity associated with complement coercion (the AMF) must involve more than simple detection of mismatching semantic properties. The latter is the standard interpretation of the N400 component. However, one of the key problems with this line of reasoning is that the M350 and N400 are merely correlates. Thus, the data of Pylkkänen et al. are fully consistent with the AMF just reflecting another subcomponent of the N400. In a subsequent study, however, Pylkkänen et al. (2009) again found an AMF effect in a different construction (the nimble climber imagined the ice survivable). In this construction, the AMF effect occurred approximately 250 ms later making it unlikely that it is a part of the N400 component. Finally, Brennan & Pylkkänen (2008a) conducted a MEG study investigating aspectual coercion in iterative semelfactives. Again, they found an AMF effect which was somewhat smaller than in the two complement coercion studies. Taken all this together, the authors conclude that the AMF reflects semantic composition in the brain.

What ERP effects are associated with complement coercion? This was investigated in two studies by Kuperberg et al. (2009) and by Choma (2006). The latter builds his three cases (18a) is derived strictly compositionally, (18b) can only be interpreted via enriched composition (Jackendoff 1996) and (18c) has no literal interpretation at all.

In Chapter 3 I formulated three hypotheses based on three classes of semantic theories on aspectual coercion. I will recapitulate them in the following.

Coercion is triggered by the occurrence of a semantic mismatch. Thus, all theories predict mismatch detection, since the coercing stimulus is semantically unexpected compared to a non-coercing control. However, for the following step – aspectual semantic repair – different theoretical options diverge on what kind of processes are hypothesized to play a role.

**Aspectual coercion involves syntactic reanalysis**

If aspectual coercion implies an update of the syntactic structure with a syntactically realized type-shifting operator (e.g. Rothstein (2004)) we expect to find neurophysiological correlates of reanalysis in aspectual coercion. This is captured by the syntactic-semantic reanalysis-hypothesis in (19).

(19) **Syntactic Reanalysis Hypothesis (SRH)**

Aspectual coercion involves (syntactic) reanalysis.

If this hypothesis is correct, coercion should lead to a P600 signature like it is typically observed in garden-path sentences. The syntactic update process is illustrated in (20). Coercion in (20c) adds an additional VP-projection compared to (20b).
The SRH also has some severe problems:

- In many languages, including German, no morpho-syntactic arguments can be provided for the SRH. Although in some cases an affix can disambiguate the Aktionsart of a verb, the same affix is also used with a variety of other Aktionsarten, too. Consider for example the affix er-. If er is added to the ambiguous verb bauen (build) we yield erbauen (edify) which unambiguously expresses an accomplishment. However, er- also appears in erlauben (admit) which is an achievement and in ermitteln (investigate) which expresses an activity. Thus, in German, Aktionsart is not realized in a morphosyntactically transparent way.
- We have already encountered differences between syntactic and aspectual processing in terms of their respective domains (see Chapter 6). This suggests that syntactic and aspectual reanalysis are different from a cognitive point of view.
- The SRH cannot explain differences between different kinds of coercion. Abstract type shift, for instance, is predicted to lead to a restructuring of the syntactic tree yielding an activity. This is similar to additive coercion, where the syntactic representation has to be updated with additional structure. Thus, both kinds of coercion should cause difficulty. But the findings of Chapter 5 show that this is not the case.

In sum, although at first glance the SRH has some attractive qualities it does not conform to the empirical findings of the preceding chapters. I will come back to it in the light of the findings of the present ERP study.

**Aspectual coercion involves lexical disambiguation**

Another class of theories assumes that a coercion operator is already structurally present (e.g. Pulman (1997); Dölling (2003a)). If no type-mismatch occurs, the operator gets interpreted as the identity function. In case of type-mismatch the appropriate operator (like add preparatory phase) has to be chosen. Coercion is therefore explained in terms of lexical ambiguity resolution. This is stated in the lexical disambiguation hypothesis in (21).

\begin{equation}
(21) \text{Lexical Disambiguation Hypothesis (LDH)}
\end{equation}

Aspectual coercion consists in the lexical disambiguation of a lexically ambiguous coercion operator.

In the evoked potentials, coercion should lead to an N400 effect due to difficulty in lexical access to the right kind of operator. I will list the pros and cons of the LDH in turn:

- The LDH allows to treat aspectual coercion in a strictly compositional fashion. But, unlike the SRH, the LDH assumes that aspectual coercion is a purely semantic phenomenon. It does not lead to structural reanalysis. Instead, the lexical meaning of a coercion operator which is structurally present right from the start has to be lexically disambiguated. The LDH is intuitively more plausible than the SRH since it reflects the semantic nature of aspectual coercion.
- The LDH offers an uniform account of both aspectual and complement coercion. In complement coercion (like the author began the book) the lexical meaning of the complement has to be interpreted in a different way than usually (e.g. not as physical object book, but as the event of writing a book). Analogously, according to the LDH, the lexical meaning of an aspectual coercion operator which is structurally already present has to be interpreted in a slightly unusual and infrequent way (that is not as the identity function).

Besides its advantages the LDH also has some problems.

- According to the LDH all sentences involve some form of aspectual coercion. By default, the phonologically empty coercion operator is disambiguated towards the
identity function. This is somewhat unsatisfying since in the majority of cases where no coercion occurs the coercion operator is superfluous.
- The LDH approaches aspectual coercion as an instance of lexical ambiguity. Thus, aspectual coercion in (22a) should involve the same cognitive processes like resolving a homonymy to a dispreferred reading in (22b).

(22) a. The mountaineer reached the summit in five hours.
   b. The bank was flooded.

Intuitively, the two examples are rather different since (22a) involves reanalysis at the sentential level while (22b) triggers reanalysis of the lexical meaning of the two homonyms of bank (reanalysis of the preferred financial institution to the dispreferred river bank).

- Moreover, disambiguated coercion operators like add preparatory phase do not directly provide the meaning of the coerced sentence. In (22a), for instance, a concrete process like climbing is needed. Therefore, the LDH implies further processing making use of world knowledge. It provides us only with the skeleton of the semantic analysis and leaves the actual resolution to pragmatics.
- Finally, there are also empirical problems concerning the LDH. The findings of Experiment 5 provide evidence that there is no processing difficulty in computing iterated readings like it is predicted by the LDH. Thus, the LDH seems to be inconsistent with the results obtained so far. However, since semelfactives occur rather often in iterative contexts the lexical frequency of iterate may be too high to lead to measurable processing difficulty during lexical disambiguation.

The LDH takes aspectual coercion to be a kind of lexical ambiguity. Since disambiguating a lexical item towards a dispreferred interpretation is reflected by the N400 component, the LDH predicts that aspectual coercion should elicit a N400 effect, too. If this prediction is correct, aspectual coercion and complement coercion give rise to the same ERP component. Such an outcome would be particularly interesting since complement and aspectual coercion can be viewed as the reverse sides of each other. Aspectual coercion leads to a semantic transformation of the context when a new lexical item is encountered which requires an argument of a different semantic type. By contrast, complement coercion involves the transformation of the new lexical item to match the context requirements. The LDH pays respect to the similarities among both coercion types.

Aspectual Coercion Makes Use of Planning
A third group of semantic theories on coercion (like Steedman (2002) and Hamm & van Lambalgen (2005)) connect aspectual coercion to the human planning capacity and predict that functional processes involved in planning tasks (like the tower of Hanoi problem) should also be active in aspectual coercion. The planning-hypothesis is stated in (23).

(23) Planning Hypothesis (PH)
Aspectual coercion involves planning and depends on finding a sequence of actions that connect the initial state with the goal state.

What does the PH predict with respect to additive coercion? An existing plan has to be updated with additional subgoals that have to be inferred via abductive reasoning. As a result, additive coercion leads to an update of the Aktionsart by enriching the semantic representation with additional propositions. In processing terms, this means that semantic content is added to working memory. In the evoked potentials this update should be reflected by a working-memory LAN in response to coerced sentences compared to non-coercing controls. Let’s have a look at the pros and cons of the PH in more detail.

- The PH offers an account of how linguistic and extralinguistic information interacts in aspectual coercion. Consider (24).

(24) The mountaineer reached the summit in five hours.

The semantics of the in-adverbial requires a process that went on for five hours. This process isn’t expressed linguistically in (24). But declarative memory contains stereotypical background knowledge about mountaineers reaching summits which is expressed in (25).

(25) HoldsAt(climb, t) ∧ HoldsAt(distance(top), t) → Happens(reach, t)

This knowledge leads to an update of the Aktionsart in working memory. Both the process (= climbing) and the incremental theme (= distance) are now introduced into working memory. Further resolution makes the initial state in (26) available.

(26) Initially(distance(0), t0)

The example illustrates that the PH provides a nice account of how the comprehender construes the meaning piece by piece using the resolution algorithm.
- The PH has been formalized using non-monotonic logic like it is needed to model abductive reasoning. An important property of these theories is that a plan can always fail due to unforeseen circumstances. This provides us with a formal theory which allows to withdraw inferences at later points in discourse.

The PH offers an essentially non-modular account of inferencing in the aspectual domain. It provides a computational framework of how temporality is encoded in
language and connects it to the notion of planning, which employs abductive reasoning about eventualities. The PH connects aspectual processing to concepts that are well established in cognitive psychology. For instance, as was shown above we can use the formalism to derive predictions about activation in different working memory systems due to additive coercion. Moreover, van Lambalgen & Stenning (2008) show that propositional logic programming can be modeled using recurrent neural nets (see Appendix A for an illustration). This makes the PH particularly interesting from a neurocognitive point of view.

The experimental design
Experiment 9 investigated ERPs elicited by sentences containing additive coercion. A sample item is provided in (28).

(28) a. Vor zwei Stunden hatte der Förster die Falle entdeckt, obwohl sie gut versteckt war.
   Two hours ago, the ranger had discovered the trap although it was hidden well.
   b. Ganze zwei Stunden hatte der Förster die Falle entdeckt, obwohl sie gut versteckt war.
   For two hours, the ranger had discovered the trap although it was hidden well.
   c. In zwei Stunden hatte der Förster die Falle entdeckt, obwohl sie gut versteckt war.
   In two hours, the ranger had discovered the trap although it was hidden well.

The design is similar to the reading time studies on additive coercion in the previous chapters. (28c) involves additive coercion of an achievement into an accomplishment. (28a) is aspectually well formed and serves as control. (28b) contains an unresolvable aspectual mismatch since an achievement expresses a punctual change of state and cannot be modified by a durative adverbial. Note that the three conditions only differ in the first word of the sentence. The critical participle and the preceding and following sentence context are identical across conditions. Note that replacing the critical participle by another verb would have yielded a different Aktionsart. This becomes immediately clear if discovered is replaced by gesucht (searched).

A second clause was added to each item always starting with although. This was done for two reasons. First, to be able to study sustained effects (e.g. a working memory LAN) some additional material is needed following the critical verb. Second, the although-clause expressed that the culminating event was unlikely to happen/hard to achieve. Thus, the second clause focuses the attention of the reader onto finding the right kind of preparatory process. This makes coercion likely to occur in the coercion condition, but, at the same time, fits well with the non-coercing controls.

Predictions
The three hypotheses predict fundamentally different ERP signatures in response to the experimental sentences.

SRH: The SRH predicts that additive coercion leads to syntactic reanalysis when the type-shift operator is integrated into the syntactic representation. This is expected to happen at the participle entdeckt (discovered) and should lead to a P600 effect in response to sentences involving coercion compared to non-coercing controls. In case of aspectual mismatch the SRH predicts essentially the same brain response since the processor tries to, but fails to update the structural representation to meet the aspectual information of the achievement verb. Accordingly, the P600 may even have a larger amplitude in the mismatch condition than in the coercion condition.

LDH: The LDH predicts difficulty in coercion due to lexical disambiguation of an ambiguous coercion operator. This should be reflected by an N400 on the critical participle entdeckt in the coercion condition compared to the control condition. The N400 component may even be of a larger amplitude in the aspectual mismatch condition than in the coercion condition since lexical processing will break down in the mismatch condition.

PH: The PH predicts that coercion leads to an update of the semantic representation with additional propositions which allow to bridge the gap between the culmination and an initial state. This should be reflected by enhanced working memory demands indicated by a working memory LAN starting at the coercing participle in the coercion condition compared to the control condition. By contrast, in the mismatch condition semantic interpretation breaks down which shouldn’t affect working memory load. The PH thus predicts no working memory LAN in the mismatch condition relative to control.

The hypotheses aren’t mutually exclusive. Thus, in the evoked potentials we may find evidence for more than one hypothesis.

Aktionsart compared to other aspects of temporality
Besides sentences investigating the brain response to the processing of aspectual coercion and aspectual mismatch, the present ERP study also included sentences with tense violations. If the latter elicited different ERP signatures than the former this would provide evidence that the processor treats these two facets of temporality qualitatively different. If we will find a smaller processing domain of tense than of Aktionsart this will provide evidence that the processor can interpret tense information before the situation type has been determined. Tense and Aktionsart will then have to be considered different temporal subsystems not only from a semantic point of view, but also from a processing perspective.
7.3.2 Method

Participants

24 German native speakers (14 female, mean age 24.2, minimum 19 years, maximum 31 years) from the University of Tübingen participated in the experiment for 22.5 Euros. None of the participants took part in the preceding experiments. All of them were classified as right handed according to the Edinburgh Handedness Inventory (Oldfield 1971). The EEG data of three additional subjects were excluded from the analysis due to excessive noise artifacts. Another subject was excluded because she judged the coercion condition not sensible at all (39 out of 40 judgments). All participants had normal or corrected vision. None had neurological or psychiatric disorders or reported neurological traumas.

Apparatus

The experiment was conducted in a dimly lit, electrically shielded and sound-attenuating booth. Humidity was kept constant at 60 percent. Participants were seated about 80 cm in front of a VGA monitor and stimuli were presented via an IBM compatible computer using E-Prime software. The responses were collected by means of a PST serial response box.

Materials

120 triples like in 28 were constructed using 40 transitive unambiguous achievement verbs each appearing in three different sentence frames. 80 items testing the processing of tense were constructed in two conditions (mismatch versus control). A sample item is provided in (29). A list of all experimental items is provided in appendix K.

(29) a. Letzte Woche bekam der fleissige Mitarbeiter die fällige Prämie. Last week received the eager employee the due bonus

b. Nächste Woche bekam der fleissige Mitarbeiter die fällige Prämie. Next week received the eager employee the due bonus

(29a) is a well formed German sentence and serves as control since the event time lies within the time frame of the adverbial. (29b) contains a tense violation. The reference time is in the future while the event time denoted by the verb lies in the past.

Experimental items testing Aktionsart and tense were constructed such that the conditions only differed in the first word. Mismatch was always induced by the main verb which was incompatible with the sentence initial adverbial either due to its Aktionsart or its tense. In the Aktionsart experiment the critical participle was the ninth segment while in the tense experiment the critical verb was the third segment. The 120 Aktionsart and 80 tense items were distributed over six lists according to a latin square design. In addition, 100 distractor items were added to each list. 30 of them were sensible and 70 didn’t make sense. Overall, half of the sentences in the experiment were sensible.

To prevent participants from making eye movements all segments were maximally 14 characters long. The frequency of the occurring words were checked using all written corpora in the Cosmas corpus. All the words used in the experiment had more than 100 occurrences.

The 300 sentences of each list were distributed over six blocks each containing 50 sentences using counterbalancing procedures such that every block contained either six or seven items in each condition. Every sentence was accompanied by a comprehension question which queried the sensibility of the corresponding sentence. The questions always appeared with two answer alternatives (yes or no). Sensible and non-sensical sentences were distributed over blocks so that the ratio of yes-answers in relation to no-answers was 1:1 across blocks.

Procedure

An experimental trial consisted of the presentation of a sentence segment by segment with a fixed presentation rate of 800 ms per segment. Immediately after the sentence had been presented a comprehension question appeared on the screen querying whether the sentence made sense. Subjects had four seconds to provide a yes/no-judgment. Sample experimental trials for Aktionsart and tense are shown in Table 7.1.

Table 7.1. Time lapse for sample trials in Experiment 9

<table>
<thead>
<tr>
<th>Segment</th>
<th>Aktionsart</th>
<th>Duration</th>
<th>Tense</th>
<th>Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fixation cross</td>
<td>+</td>
<td>1000</td>
<td>+</td>
<td>1000</td>
</tr>
<tr>
<td>1. Segment</td>
<td>Ganze</td>
<td>800</td>
<td>Letzte</td>
<td>800</td>
</tr>
<tr>
<td>2. Segment</td>
<td>zwei</td>
<td>800</td>
<td>Woche</td>
<td>800</td>
</tr>
<tr>
<td>3. Segment</td>
<td>Stunden</td>
<td>800</td>
<td>bekam</td>
<td>800</td>
</tr>
<tr>
<td>4. Segment</td>
<td>hatte</td>
<td>800</td>
<td>der</td>
<td>800</td>
</tr>
<tr>
<td>5. Segment</td>
<td>der</td>
<td>800</td>
<td>fällig</td>
<td>800</td>
</tr>
<tr>
<td>6. Segment</td>
<td>Rentner</td>
<td>800</td>
<td>Mitarbeiter</td>
<td>800</td>
</tr>
<tr>
<td>7. Segment</td>
<td>den</td>
<td>800</td>
<td>die</td>
<td>800</td>
</tr>
<tr>
<td>8. Segment</td>
<td>Schlüssel</td>
<td>800</td>
<td>fällig</td>
<td>800</td>
</tr>
<tr>
<td>9. Segment</td>
<td>gefunden</td>
<td>800</td>
<td>Prämie</td>
<td>800</td>
</tr>
<tr>
<td>10. Segment</td>
<td>obwohl</td>
<td>800</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. Segment</td>
<td>er</td>
<td>800</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12. Segment</td>
<td>in der</td>
<td>800</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13. Segment</td>
<td>Schublade</td>
<td>800</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14. Segment</td>
<td>verstieckt war</td>
<td>800</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Question</td>
<td>makes sense?</td>
<td>max. 4000</td>
<td>makes sense?</td>
<td>max. 4000</td>
</tr>
<tr>
<td>Inter trial pause</td>
<td>empty screen</td>
<td>1700</td>
<td>empty screen</td>
<td>1700</td>
</tr>
</tbody>
</table>

Note: duration in milliseconds.

7. Assuming that coercion was judged to be sensible.
Each segment was presented individually in the center of the screen in yellow letters on a blue background. Size of font was chosen in a way that no segment was bigger than $5^\circ \times 0.5$ degrees of the visual angle of a subject. If a subject did not answer within the time limit the next trial started automatically. Subjects had to respond by pressing a button either with the left or with the right index finger. Fingers were uniquely mapped onto the “yes” and the “no” judgment. For half of the subjects the “yes” judgment was mapped to the right index finger while the mapping was reversed for the other half. The participants were instructed to provide their judgments as fast as possible. They were told not to move or blink while a sentence was displayed.

After reading a written instruction they received a practice block block with ten practice items which contained no aspectual violations. During practice they received explicit feedback about the errors they made. The practice session was followed by six experimental blocks. After each block there was a pause to give subjects the opportunity to relax. The whole experiment (including the application of electrodes) lasted for approximately three hours. The experimental session took approximately 75 minutes.

EEG recordings
The EEG-activity was measured with 30 Ag/AgCl-electrodes which were attached to the scalp using the Easycap system at FP1, FP2, F7, F3, Fz, F4, F8, FC5, FC1, FC2, FC4, T7, T3, Cz, C4, T8, the left mastoid, CP5, CP1, CP2, CP4, P7, P3, Pz, P4, P8, O1, Oz and O2. Electrode positions were chosen in accordance with the international 10/20 system (Jasper 1958).

Each electrode was referenced to the right mastoid and the ground electrode was positioned at FPz. Horizontal (HEOG) and vertical (VEOG) eye activity was measured by placing two electrodes 2 cm lateral to the right and the left canthus for the HEOG and two electrodes 3 cm above and below the pupil of the right eye for the VEOG. All electrophysiological signals were digitized with a frequency of 500 Hz. The impedances of each electrode were kept below 3 kΩ for the EEG electrodes and below 5 kΩ for the EOG electrodes.

Data analysis of the EEG data: Aktionsart
The continuously recorded electrophysiological data were rereferenced to linked mastoids. Then they were epoched for every single trial by cutting segments which started 200 ms before the onset of the first word and ended with the end of the sentence. All epochs were filtered by a 20 Hz low-pass-filter (with a filter slope of 12 dB per octave). They were then further divided. For the investigation of effects at the participle segments were cut which began 200 ms before its onset and lasted until 500 ms after the end of the sentence. The 200 ms before the presentation of the participle were used for baseline correction. Similarly, segmentation was accomplished for the beginning of the sentence, cutting segments which started 200 ms before the onset of the first word and ended with the onset of the participle.

All trials contaminated by blinks, eye movement artifacts, or other electrical noise were rejected during data analysis.

Data analysis of the EEG data: Tense
The rereferenced EEG data were epoched by cutting segments which started 200 ms before the onset of the verb and lasted 1500 ms post stimulus. The 200 ms before the presentation of the verb were used for baseline correction. Trials contaminated by artefacts were excluded from further analysis.

Statistical analysis
Repeated measures ANOVAs were computed using two regions of interest (ROIs) and mean amplitude as dependent variable. The time interval of 300 to 500 ms post onset served to investigate a morphosyntactic LAN effect or an N400 effect. The time window from 500 to 900 ms post onset was used to investigate P600 effects and/or a working memory LAN. ROIs were identical for aspectual conditions and for tense conditions. ANOVAs included the factors aspect and tense respectively (three levels in the aspectual cases, two levels in the tense conditions) and electrode (29 levels).

To investigate the topographical characteristics of ERP effects, four topographical regions were defined: left anterior electrode sites (containing FP1, F7, F3, FC5 and FC1), right anterior sites (FP2, F4, F8, FC2 and FC6), left posterior sites (CP5, CP1, P7, P3 and O1) and right posterior sites (CP2, CP6, P4, P8 and O2). The mean amplitudes of these four regions were used to compute ANOVAs including the factors anteriority and hemisphere.

Further, ANOVAs were computed analyzing behavioral data (judgments and judgment times). Planned comparisons were computed using paired t-tests always applying a Bonferroni corrected a.

The grand averages shown in the Figures are low pass filtered with a 10 Hz filter. The statistical analyses, however, were computed on the basis of unfiltered data. Degrees of freedom were always corrected using the Greenhouse-Geisser correction (Huynh 1978). In the text the uncorrected degrees of freedom are reported.
7.3.3 Results

Behavioral data

Figure 7.2 shows the percent "yes, makes sense" answers across conditions.

![Figure 7.2. Semantic acceptability judgments in the ERP study](image)

In the aspectual cases participants judged aspectual coercion as sensible in 75% compared to 87% in the control condition and 15% in the mismatch condition. Thus, although coercion was judged nonsensical more often than control, subjects chose a coerced interpretation in the majority of the cases. An ANOVA yielded a significant main effect of aspect ($F (2, 46) = 354.53; p < .01$). A paired t-test revealed that the control condition was judged sensible more often than coercion ($t (23) = 3.85; p < .01$). Judgment times (mismatch: 752 ms, coercion: 808 ms, control: 818 ms), however, did not differ across conditions ($F (2, 46) = 1.87; p = .16$).

Sentences containing a tense violation only received 11 percent "yes, makes sense" responses while control was judged sensible in 97 percent. The answer times were slightly faster in the mismatch condition (640 ms) than in the control condition (694 ms). This difference was significant ($t (23) = 2.14; p < .05$).

ERP data: Aspect

Grand averages of the aspectual conditions on the participle are depicted in Figure 7.3.

![Figure 7.3. Grand averages elicited by aspectual mismatch, coercion and control. Zero marks the onset of the participle](image)
The distribution of the working memory LAN elicited by coercion is illustrated in Figure 7.5 showing the mean difference between coercion and control.

In the global ANOVA, both the interaction between aspect and electrode \((F(56, 1288) = 4.04; p < .01)\) and the main effect of aspect \((F(2, 46) = 5.83; p < .01)\) were significant. To break these effects further down, two ANOVAs with the within factors anteriority, hemisphere and aspect were computed.\(^8\) The first compared mismatch to control and the second coercion to control. In the former analysis there was both a significant interaction between aspect (mismatch vs. control) and anteriority \((F(1, 23) = 7.62; p < .01)\) resulting from the more posterior positivity. The latter analysis also revealed both a significant interaction between aspect (coercion vs. control) and anteriority \((F(1, 23) = 6.28; p < .05)\) due to the anterior negativity. No other effects involving the factor aspect were significant.

Figure 7.6 shows the mean amplitudes of anterior and posterior electrode sites with respect to both hemispheres.

To further investigate the sustained negativity the time window from 900 ms to 1500 ms was analyzed. An ANOVA with the within factors anteriority, hemisphere and aspect (coercion vs. control) resulted in a significant interaction between aspect and anteriority \((F(1, 23) = 8.70; p < .01)\). Besides a main effect of hemisphere no other effects were significant. This interaction is due to the negativity extending right to the following word which started 800 ms post onset. Figure 7.2 shows the mean amplitude of the four topographical regions in the 900 to 1500 ms time window. The mismatch condition could not be analyzed in this ROI because of the preceding P600 effect.

**ERP data: Tense**

Grand averages of tense violations compared to controls are depicted in Figure 7.7. Tense mismatch compared to control elicited a positive deflection in the ERP waveform. This positivity started about 500 ms after stimulus presentation. Like in the case of

---

\(^8\) The mean of left anterior electrodes was computed by averaging over electrodes FP1, F7, F3, FC5 and FC1. Similarly, FP2, F4, F8, FC2 and FC6 were grouped together into right anterior electrode positions; CP5, CP1, P7, P3 and O1 were left posterior and CP2, CP6, P4, P8 and O2 were right posterior positions.

---
aspectual mismatch the positivity was strongest at central and parietal midline electrode sites. The distribution is depicted in Figure 7.8 which shows the mean difference scores between tense violations and controls.

An ANOVA for the 300–500 ms time window yielded no significant effect involving tense. In the 500 to 900 ms ROI an ANOVA revealed a significant interaction between tense and electrode \( (F(28, 644) = 3.69; p < .01) \) and a significant main effect of tense \( (F(1, 23) = 9.36; p < .01) \) reflecting the positive shift which was most pronounced at central and parietal midline electrode sites.

### 7.3.4 Discussion

Acceptance rates clearly showed that participants were paying attention to Aktionssart and tense. While sentences with an aspectual or a tense mismatch were judged as semantically ill-formed, subjects accepted both aspectual coercion and control sentences. That coercion sentences were rejected slightly more often than aspectual controls indicates that they were more difficult to understand. Equal question answer times in additive coercion and aspectual control further demonstrate that participants already had computed a coerced interpretation when they finished reading a sentence.

In the evoked potentials a P600 effect was elicited by participles that introduced an unresolvable aspectual mismatch. By contrast, aspectual coercion did not elicit a P600, but was reflected by a sustained working memory LAN which had the biggest amplitude way into the subsequent word. The qualitatively different ERP effects in the two conditions are interesting since on the basis of the general model from Chapter 3 a completely different pattern of results was expected. The coercion model contained a revision phase in which the literal interpretation...
The Processing of Events

leads to an aspectual mismatch and a second phase in which coercion is carried out. This means that during mismatch detection coercion and mismatch should initially elicit the same brain responses, but only at a later processing stage coercion should differ from mismatch due to aspectual repair. The findings of the ERP study challenge this view and provide evidence for a smooth update process in aspectual reanalysis instead. We only found indication of enhanced working memory load in coercion, but crucially no evidence for aspectual mismatch detection. This pattern of results is fully compatible with a theory which assumes that in additive coercion an additional process is added to the existing representation without revising it first. A computation along these lines was already spelled out within the framework of Hamm & van Lambalgen (2005) in the derivations of Chapter 3. In order to coerce an achievement into an accomplishment the achievement scenario has to be updated by simply adding program clauses which encode a preparatory process. The culmination and the resultant state are not affected by this operation.

The experimental findings support the planning hypothesis

The working memory LAN provides support for the PH which states that coercion makes use planning. Planning sets the culmination as a goal which has to be achieved by means of an appropriate preparatory process. This process has to be inferred on the basis of world knowledge. The enrichment of the situation model is expected to enhance working memory load like it was found in our study.

By contrast, the results are not consistent either with the SRH or with the LDH. The SRH predicted a P600 effect due to coercion, but in the experiment a P600 was only found in the mismatch condition. Thus, it looks as if the meaning did not have to be revised in coercion, but was only smoothly updated with yet missing information. The LDH predicted a N400 effect in aspectual coercion. But all three aspectual conditions did not differ in the N400 ROI. Thus, the results provide evidence that coercion doesn’t involve the disambiguation of a covert coercion operator.

Aspectual coercion not leading to an N400 effect is especially interesting since this calls for a functional difference between aspectual coercion and complement coercion. While the latter involves lexical processing and leads to an N400 effect (Kuperberg et al. 2009), aspectual coercion does not seem to be resolved at the lexical level.

Processing different facets of temporality

Tense violations elicited a P600 effect at the finite verb. The mismatch effect occurred before the verb had received its arguments. This is particularly interesting since it implies that the processing domain of tense is smaller than that of Aktionsart. Tense and Aktionsart have to be considered independent temporal subsystems not only linguistically, but also cognitively.
Despite the difference in incremental size between tense and Aktionsart, mismatches in tense and Aktionsart qualitatively lead to the same brain responses. Both elicited late positive shifts which are commonly assumed to reflect difficulty in syntactic processing. However, in the present experiment the P600 was rather indexing semantic interpretation breakdown at the sentential level.

Let's first have a closer look at the tense violations. Disregarding the lexical information of the adverbial, the verb was morphosyntactically of the correct form. A mismatch occurred because the tense information was not included in the reference time introduced by the preceding adverbial. This is clearly a semantic violation.

Is it possible that readers implicitly carried out morphosyntactic repair reconciling the mismatch? I think that this is unlikely. In the experimental sentences implicit repair would mean that readers reanalyzed a sentence like (30a) into its future tense form (30b). This implies not just fixing the morphosyntax of the verb, but a complete restructuring of the sentence.

Moreover, in aspeccual mismatch there is not even the possibility to implicitly repair the sentence on morphosyntactic grounds. We are thus dealing with a semantic violation that isn't reflected in the syntax. The positive shift due to aspeccual mismatch provides strong evidence that the P600 is not limited to syntactic integration difficulty, but more generally indexes interpretation difficulty at a supraparalexical level.

The incremental computation of a temporal model
The EC system can be used to model the findings of the ERP study taking Aktionsart and tense into consideration simultaneously. To illustrate that I will derive the EC representation of (31).

Gestern hatte Hans den Gipfel in drei Stunden erreicht.

Yesterday had Hans the top in three hours reached

When the reader encounters the first phrase yesterday it translates into the integrity constraint (37a) and the scenario (37b).

The Processing of Events

Fulfilling the instruction expressed by the integrity constraint (32a) leads to the construction of a minimal temporal model including the utterance time now and the reference time R which lies at most 24 hours in the past of the utterance time. Next, the reader encounters hatte which introduces an additional constraint expressing the past perfect (33).

In (33) f refers to a state which comes about by some event e happening at the event time E which lies in the past of R. This means that the past perfect automatically introduces a yet abstract achievement formula into the scenario \( \text{Initiates}(e, f, E) \). The event e and and the state fluent f have to be unified with a concrete event and a state further downstream the sentence. The integrity constraint (33) serves as an instruction to update the temporal model with the event time E which lies in the past of the reference time R.

Note that a tense mismatch (e.g. *yesterday will ...*) would have resulted in an integrity constraint expressing a contradiction (e.g. ? ... R < now, R > now ...). Since a contradictory constraint cannot be fulfilled in any model, semantic processing breaks down.

The next two phrases Hans and den Gipfel introduce two new discourse referents into the model. The information up to the direct object is expressed in (34). Introducing new discourse referents does not change any of the temporal properties of the model.

The in-adverbial calls for two temporal entities: an event and a three hour long process which is finished when the event happens. Thus, we are now dealing with a yet abstract accomplishment scenario (35b). The according integrity constraint is stated in (35a).

Chapter 7. The processing of temporality in the brain

2nd proofs
The findings of the EEG experiment complement these data. The evoked potentials were arguably no issue in aspectual mismatch. P600 to be an index of syntactic integration difficulty and syntactic reanalysis which were also sustained. Interestingly, aspectual mismatch lead to a qualitatively different brain response, namely a P600 effect which wasn’t present in additive coercion. This pattern of results is only compatible with theories of coercion which allow for a smooth update of the situation model without revisiting it first.

The EEG study also included tense violations occurring at a point where the finite verb had not received any arguments. Similarly to aspectual mismatch tense violations also elicited a P600 effect, but the effect was already present at the bare verb. This is interesting since it suggests different processing domains of tense and Aktionsart. During online interpretation tense can be computed independently from aspect. The temporal model is not computed as a whole, but is constructed piece by piece as the reader moves along the sentence.

Finally, it is rather surprising that both tense violations and aspectual mismatch lead to a P600 in the evoked potentials. This kind of brain response has traditionally been associated with syntactic processing while semantic processing has been taken to be reflected by the N400. In the present study the P600 was indicating the breakdown of semantic interpretation, but the tested constructions did not differ in terms of the N400. These findings suggest a more liberal interpretation of the P600 as an index of interpretation difficulty at the phrasal level.

How do the results of the present EEG study relate to other psycholinguistic investigations on coercion phenomena? Interestingly, in their ERP studies on complement coercion Choma (2006) and Kuperberg et al. (2009) reported N400 effects in the coercion condition. The difference in ERP signatures provides evidence that the two kinds of coercion are functionally different. While complement coercion requires lexical processing, that is shifting a lexical item from one type to another, aspectual coercion involves supralexical processing, the transformation of a complex situation model into another type of situation.

It is further worth noting that the findings of the present study can be made compatible with the neurolinguistic model of Hagoort (2003b). In Hagoort’s model unification failure elicits a P600. Mismatches of Aktionsart and tense can also be viewed as involving unification failure despite the fact that in the processing of tense and aspect we are dealing with unification of constants and variables in the semantic representation. In contrast, aspectual coercion did not involve binding problems. Instead a fluent variable had to be unified with a process which wasn’t expressed linguistically, but had to be inferred from world knowledge. Therefore, no P600 is expected given Hagoort’s interpretation of the late positive shift.

The data are difficult to integrate into Friederici’s neurocognitive model of sentence processing (Friederici 1999; Friederici 2002). This is due to the fact that she reserves the P600 to be an index of syntactic integration difficulty and syntactic reanalysis which were arguably no issue in aspectual mismatch.

Next the verbal information of erreicht is integrated into the scenario via unification: e is unified with the reach event and fstate(e) is unified with the state be on top. Although unification allows to replace the event variable e and the state fluent f with concrete eventualities, the achievement verb includes no information that allows to instantiate the process variable with a constant. To achieve this goal additional information is needed which has to be inferred on the basis of world knowledge. Holding this additional information in working memory leads to increased memory load. The update proceeds in smooth fashion. It leaves the scenario intact and just adds information to it. This explains why in additive coercion the ERP signature showed no indication of a revision.

What happens in case of an aspectual mismatch? Instead of an in-adverbial a for-adverbial preceded the participle. Encountering a for-adverbial introduces an activity scenario including a process variable only. Neither the reach event nor the resultant state can be unified with this variable and unification failure causes a breakdown of semantic interpretation.

### 7.4 Summary and conclusions

In this chapter I have presented electrophysiological data to elucidate on the processing of Aktionsart and tense. Three points about the study are especially interesting.

The reading time studies of the preceding chapters provided evidence that additive coercion is causing processing difficulty. In these experiments additive coercion lead to longer reading times. This effect was sustained until the end of the sentence. The findings of the EEG experiment complement these data. The evoked potentials indicated enhanced working memory demands caused by additive coercion which were also sustained. Interestingly, aspectual mismatch lead to a qualitatively different brain response, namely a P600 effect which wasn’t present in additive coercion. This pattern of results is only compatible with theories of coercion which allow for a smooth update of the situation model without revisiting it first.

The EEG study also included tense violations occurring at a point where the finite verb had not received any arguments. Similarly to aspectual mismatch tense violations also elicited a P600 effect, but the effect was already present at the bare verb. This is interesting since it suggests different processing domains of tense and Aktionsart. During online interpretation tense can be computed independently from aspect. The temporal model is not computed as a whole, but is constructed piece by piece as the reader moves along the sentence.

Finally, it is rather surprising that both tense violations and aspectual mismatch lead to a P600 in the evoked potentials. This kind of brain response has traditionally been associated with syntactic processing while semantic processing has been taken to be reflected by the N400. In the present study the P600 was indicating the breakdown of semantic interpretation, but the tested constructions did not differ in terms of the N400. These findings suggest a more liberal interpretation of the P600 as an index of interpretation difficulty at the phrasal level.

How do the results of the present EEG study relate to other psycholinguistic investigations on coercion phenomena? Interestingly, in their ERP studies on complement coercion Choma (2006) and Kuperberg et al. (2009) reported N400 effects in the coercion condition. The difference in ERP signatures provides evidence that the two kinds of coercion are functionally different. While complement coercion requires lexical processing, that is shifting a lexical item from one type to another, aspectual coercion involves supralexical processing, the transformation of a complex situation model into another type of situation.

It is further worth noting that the findings of the present study can be made compatible with the neurolinguistic model of Hagoort (2003b). In Hagoort’s model unification failure elicits a P600. Mismatches of Aktionsart and tense can also be viewed as involving unification failure despite the fact that in the processing of tense and aspect we are dealing with unification of constants and variables in the semantic representation. In contrast, aspectual coercion did not involve binding problems. Instead a fluent variable had to be unified with a process which wasn’t expressed linguistically, but had to be inferred from world knowledge. Therefore, no P600 is expected given Hagoort’s interpretation of the late positive shift.

The data are difficult to integrate into Friederici’s neurocognitive model of sentence processing (Friederici 1999; Friederici 2002). This is due to the fact that she reserves the P600 to be an index of syntactic integration difficulty and syntactic reanalysis which were arguably no issue in aspectual mismatch.
7.4.1 A question for future research

Unfortunately, EEG data are not well-suited to localize the effects of mismatch and coercion. For this purpose, a functional imaging technique like fMRI would be much more appropriate (for an introduction to the fMRI methodology see Huettel, et al. (2004)). On the basis of the findings of the ERP study it is, however, possible to derive predictions about the locus of effect to be expected in an imaging experiment. The Immediacy Model of Hagoort (2003b) and the more recent MUC (memory, unification, control) model by Hagoort (2005) make interesting neurolinguistic predictions about brain areas which may play a role.

Aspectual mismatch was characterized by unification failure. The for-adverbial is in need of a process, but the achievement only consists of an event plus a state. Thus, the achievement cannot be unified with the adverbial. This kind of unification failure is purely semantic in nature. Hagoort (2005) identifies the left inferior frontal gyrus which encompasses Brodmann areas 6, 44, 45 and 47 (LIFG) with the unification space in which the phonological, syntactic and semantic information of lexical items are bound together into larger structures. Interestingly, different kinds of information are claimed to be processed in specialized subregions of LIFG. According to the model, semantic unification is carried out in BA 47/45, while syntactic and phonological information are brought together in more posterior-ventral parts of the prefrontal cortex (syntax: BA 45/44 vs. phonology: BA 44/6). Under the assumption that aspectual mismatch resulted in a failure in semantic binding, it should lead to enhanced activation in BA 47/45 of LIFG.

In order to resolve additive coercion the reader has to infer a missing piece of information and unify this information with the linguistic representation. He has to come up with a plan that allows to reach the goal state via a preparatory process. This involves two components. First, the culmination has to be set as a goal and second, declarative memory has to be searched to retrieve a suitable process. The former should lead to activation of control (a network involving the anterior cingulate cortex and the dorsolateral prefrontal cortex, BA46/9). The latter is the memory module in Hagoort's (2005) model and is located in the left temporal cortex.

Interestingly, the topographical region of control is intimately connected to planning as is shown by lesion studies (e.g. Shallice & Burgess (1991)). Even if a typical patient suffering from a lesion of prefrontal cortex has all the declarative knowledge at his disposal that is needed to make a complex plan he is often unable to integrate the parts into a coherent whole. For instance, in their study on the water jug task – a task that requires the subject to move from an initial state to a goal state via a number of intermediate steps – Colvin, et al. (2001) found that patients with a lesion of the left dorsolateral prefrontal cortex had difficulty with that task: they seemed to be unable to appropriately inhibit a response in keeping with the final goal. Similarly, brain imaging studies on the Wisconsin card sorting test (= "the neuropsychological test" for prefrontal lesions) show that the prefrontal cortex is highly active in this task, especially when a new rule has to be adapted (Monchi, et al. 2001). Considering this evidence in the light of EC it would be interesting to see whether aspectual coercion involves the same brain areas.

To sum up, based on Hagoort's (2005) model, a double dissociation of aspectual mismatch and additive coercion is expected. While mismatch should lead to enhanced activity in the anterior part of the LIFG, additive coercion should lead to enhanced activity in control regions (dorsolateral parts of the prefrontal cortex) and in memory (left temporal cortex). Testing these predictions must be left for future research.
CHAPTER 8

The specified processing model of aspectual reanalysis

I started with a general model of aspectual reanalysis and formulated hypotheses about different aspects of the model. Based on the experimental findings of the previous chapters the model can now be fully specified. Table 8.1 lists the hypotheses that have been put to test.

The first three hypotheses formulated theoretical alternatives on the online processing of (ambiguous) aspectual information. Experiments 2 and 3 identified two factors that guide aspectual access. The findings of Experiment 2 show that frequency information is used to determine the most probable Aktionsart of aspectually ambiguous verbs. Experiment 3 showed that lexical frequency information can be overridden by context information from the preceding discourse: an activity verb receives an accomplishment interpretation when the preceding discourse contains an appropriate incremental theme. Taken together, the experiments indicate that aspectual processing uses the same kinds of information as syntactic processing. Furthermore, aspectual processing doesn’t take into consideration only the current sentence, but operates on the discourse representation.

The second set of hypotheses concerned difficulty in three different types of coercion. In Experiments 4a, 5 and 6 I investigated additive coercion, subtractive coercion and abstract type shift, respectively. Only additive coercion lead to measurable disruption during reading. Additive coercion may be particularly difficult because the situation model has to be enriched with eventualities that are not encoded linguistically. This makes it the aspectual analogue of complement coercion. With respect to the other two types of aspectual coercion, viz. subtractive coercion and abstract type-shift, we have to be careful in interpreting the results because of the danger of type II errors. In the next section I will describe one last experiment which was conducted to replicate the results from these experiments.

The third set of hypotheses concerned the processing domain in aspectual interpretation. The results of Experiments 4b, 7 and 8 indicate that only complete verb-argument structures are aspectually specified. Modifying a verb and its direct object by an aspectually coercing or mismatching adverbial didn’t lead to immediate processing difficulty, instead difficulty was delayed until the extraposed subject.
Table 8.1. Hypotheses that have been confirmed (= ✓) or disconfirmed (= No)

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>✓</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Determining Aspect:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Probabilistic Parsing</td>
<td>✓</td>
<td>No</td>
</tr>
<tr>
<td>Lazy Parsing</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>Aspectual Enrichment</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>Types of Coercion:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Abstract Type Shift</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>Non-Monotonic Update</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>Aspectual Enrichment</td>
<td>✓</td>
<td>No</td>
</tr>
<tr>
<td>Processing Domain:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Late Aspectual Interpretation</td>
<td>✓</td>
<td>No</td>
</tr>
<tr>
<td>Complete VP</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>Incremental Aspectual Interpretation</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>Cognitive Processes:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Syntactic Reanalysis</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>Lexical Disambiguation</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>Planning</td>
<td>✓</td>
<td>No</td>
</tr>
</tbody>
</table>

Similarly, in Experiment 4b there was no processing difficulty at the adverbal phrase when it intervened between the verb and the direct object. The completion study (Experiment 7) showed that the lack of early effects isn’t due to readers predicting meaningful continuations of yet incomplete minimal sentences. Hence, incomplete verb-argument structures don’t seem to be aspectually specified yet.

Finally, Experiment 9 investigated the processing of additive coercion using event related potentials. I formulated three hypotheses about the processes that might be involved in aspectual coercion. The first hypothesis, syntactic reanalysis, stated that additive coercion involves adding structure to the phrase marker in order to decompose the coerced eventuality syntactically. Restructuring the syntactic tree should lead to a syntax-related ERP component such as a morphosyntactic LAN or a P600, but additive coercion didn’t elicit any of these effects. The second hypothesis, lexical disambiguation, stated that coercion involves disambiguating a covert coercion operator (which is usually interpreted as the identity function \( \lambda P \)P) towards an infrequent reading (e.g. Egg (2005)’s add preparation). This should be reflected by difficulty in lexical processing. In Experiment 9, however, we didn’t find any indication of a N400 effect. This makes it unlikely that aspectual coercion is resolved at the lexical level. The last hypothesis, planning, stated that in additive coercion an existing plan is enlarged with additional subgoals and that enlarging the plan should enhance working memory load. In Experiment 9, a working memory LAN was found confirming this prediction. The findings of Experiment 9 speak against a revise and reanalyze account of additive coercion, but go together with a smooth update of Aktionsart. This is because there was a double dissociation between aspectual coercion and aspectual mismatch.

The possibility of coercion is limited, however. Throughout the dissertation I compared coercion to aspectual mismatch which was clearly judged as nonsensical. The mismatch cases – achievements modified by a for-adverbial – don’t allow aspectual coercion. Where should the line be drawn between sentences that can be coerced and those that don’t? The mismatch sentences become acceptable when turned into the progressive. Consider the contrast between (1a) and (1b).

(1) a. *John reached the top for an hour.
   b. John was reaching the top for an hour.

The difference between (1a) and (1b) lies in the number of coercions that are required at once. In (1a) the adverbial forces two coercions simultaneously: additive coercion turning the achievement into an accomplishment and subtractive coercion picking out the preparation. This is too much to be carried out in a single composition step. In (1b), however, the progressive already shifts the achievement into an accomplishment. Given that only one step, subtractive coercion, is needed to fit the requirements of for. Thus, in contrast to Moens & Steedman (1988) and Pulman (1997) simultaneous recursion over coercions doesn’t seem to be possible.

Putting everything together, what does the general picture of aspectual processing look like? During parsing aspect the processor exploits probabilistic information to determine which Aktionsart is most likely. The processing unit relevant for aspectual composition is rather big compared to the word-by-word processing typically assumed in syntax; instead of single words it’s whole verb-argument structures that are required to determine aspectual properties. During normal processing the aspectual type of the new input fits the type of the existing representation and can be easily integrated. In aspectual coercion, however, the new input shifts the existing representation into an Aktionsart of a different type. In one subtype of aspectual coercion, subtractive...
coercion, this will lead to a non-monotonic update of the situation model. Inferences that were valid before do not carry over to the new representation. However, subtractive coercion doesn’t seem to cause difficulty in processing. The second type of coercion, abstract type shift, transforms a complete event nucleus into a different situation type, for instance by iterating it. This type of coercion doesn’t seem to be difficult, either. Finally, in additive coercion, the existing representation doesn’t contain enough information to match the input requirements of an item carrying aspectual information. Inferring the right kind of eventuality enhances processing load. But if the linguistic context contains the missing piece of information, processing difficulty is eliminated. Aspectual processing has access to the discourse representation and makes direct use of contextual information.

As can be seen, all parts of the general model of aspectual processing from Chapter 3 have been addressed. The experimental findings have put us in a position to decide between theoretical alternatives and to further refine the model.

8.1 A replication of Experiments 5 and 6: Experiment 10a/b

Two conclusions from the preceding section stand on shaky ground. I have argued that neither subtractive coercion nor abstract type-shift is taxing in online comprehension. This was based on an interpretation of null effects between the coercion condition and the control condition in Experiments 5 and 6. This line of reasoning is problematic, however, because of the danger of type II errors, that is, falsely adhering to the null hypothesis although there really is a difference. One way to make sure that this was not the case is try to replicate the results and see if they come out the same. This was done in the context of a larger experiment (Experiment 10a/b) which included the right kind of conditions to replicate the earlier results. Above that, the experiment tested the processing consequences of iterative coercion in two aspectual classes, semelfactives and accomplishments (the study was also published in Bott (2008)).

Let’s look at subtractive coercion in accomplishments and iterative semelfactives again.

(2) a. Der Bauarbeiter belud die Schubkarre eine Stunde lang. The construction worker loaded the wheelbarrow for an hour. 
   b. Der Bauarbeiter belud die Schubkarre in einer Stunde. The construction worker loaded the wheelbarrow in an hour.

Sentence (2a) involves subtractive coercion of an accomplishment into an activity, (2b) is the control condition. The time interval expressed by the for-adverbial in (3) is too long to allow for a single loading event. Thus, the accomplishment is iterated into a series of loading events. This is an instance of a habitual reading.

(3) Der Bauarbeiter belud die Schubkarre zwanzig Jahre lang. The construction worker loaded the wheelbarrow for twenty years.

Since the previous experiments showed that a mismatch condition is a useful thing to have for comparison with aspectual coercion I included one in this experiment, too. Combining accomplishments with long in-adverbials yields implausible readings, a pragmatic mismatch condition illustrated in (4).

(4) Der Bauarbeiter belud die Schubkarre in zwanzig Jahren. The construction worker loaded the wheelbarrow in twenty years.

The accomplishments were compared with coerced and non-coerced semelfactives. (5a) is an iterated semelfactive and (5b) the relevant control condition.

(5) a. Der Junge nieste zwanzig Minuten lang. The boy sneezed for twenty minutes.
   b. Der Junge nieste vor zwanzig Minuten. The boy sneezed twenty minutes ago.

The experiment contained two sets of materials. The first, Experiment 10a, encompassed accomplishments modified by for- and in-adverbials. Additionally, the factor length of time (short adverbial vs. long adverbial) was manipulated in a two-by-two factorial design. The second, Experiment 10b, contained iterative semelfactives and control sentences denoting a single event.

8.1.1 Materials

40 sets of accomplishment sentences were constructed like the sample item in (6). Appendix M contains the complete set of items.

(6) a. Der Bauarbeiter belud die Schubkarre eine Stunde lang. The construction worker loaded the wheelbarrow for an hour.
   b. Der Bauarbeiter belud die Schubkarre in einer Stunde. The construction worker loaded the wheelbarrow in an hour.
   c. Der Bauarbeiter belud die Schubkarre zwanzig Jahre lang. The construction worker loaded the wheelbarrow for twenty years.
   d. Der Bauarbeiter belud die Schubkarre in zwanzig Jahren. The construction worker loaded the wheelbarrow in twenty years.

... dann wurde er woanders eingeteilt. ... then was he somewhere else assigned
   ... then he was assigned to another duty.
   ... dann wurde er in Ruhe gesendet. ... then he was retired.
The short and long adverbials were matched in length. Short adverbials had a mean of 17.6 characters, long adverbials a mean of 17.4 characters. A two-samples t-test revealed that the numerical difference was far from significant (t(158) = .67, p = .51).

The experimental sentences were always followed by a then-clause. There were two types of then-clauses depending on the adverbial. Following short adverbials the second clause introduced a concrete event which can plausibly interrupt the preparation. Participants are thus expected to understand subtractive coercion in (6a) as implying that the culmination didn’t happen. By contrast, in (6b) the event introduced by then is expected to be interpreted as abutting the culmination. Following long adverbials the event was rather abstract and shifted the focus to higher order events. This was done to make it coherent with the habitual interpretation of the iterative accomplishment. In the above example it situates the event at the level of life episodes like education, job, retirement. Crucially, the two continuations only differ after the first three words of the second clause. Notice that only the processing of the first sentence will be relevant in the reading time study.

In addition, 20 sets of semelfactive sentences like the sample item in (7) were constructed. The complete set of items can be found in Appendix N.

(7) a. Den ganzen Morgen niesste der Junge laut im Klassenzimmer,…  
The whole morning sneezed the boy loudly in the classroom,…  
… dann entschuldigte er sich bei seinen Mitschülern.  
… then apologized he to his classmates.  
… then he apologized to his classmates about it.

The adverbial was presented in topicalized position as in the study by Brennan & Pylkkänen 2008a. This was done to be able to compare the experimental findings of the two studies. A then-clause was added to the semelfactives to make them maximally similar to the accomplishments.

8.1.2 Pretesting the readings

Before turning to the reading time experiment it is crucial to establish whether the above sentence types have the intended meanings. To this aim, I tested the readings in a pretest.

Methods

The pretest was a questionnaire study that was administered over the internet using Webexp 2 (Mayo et al. 2006). For each sentence participants (N = 24) had to provide two judgments. First, they had to decide whether the sentence expressed that the (with accomplishments: culminating) event happened and if so, whether it happened only once or repeatedly.

Only the first three conditions (6a–c) of the accomplishments were included in the pretest because only these should receive a sensible interpretation. From the 40 accomplishment sentences and the 20 semelfactive sentences, six lists were created according to a latin square design. The same set of 30 fillers were added to each list resulting in a total of 70 items. Each list was individually pseudorandomized in two orders. The pretest took approximately 20 minutes.

Results and discussion

The results are depicted in Figure 8.1. Accomplishments modified by short in-adverbials and long for-adverbials overwhelmingly received “yes, the culminating event happened” responses (short in: 76.6%; long for: 81.8%), but the short for-adverbials received “yes” responses only 16.4% of the time. ANOVAs with the within-factor adverbial (two levels: short for vs. short in) comparing the “yes”-responses in the two constructions with short adverbials revealed a reliable difference between short for and short in adverbials (F(1,23) = 107.18 MSE = 4.35, p < .01; F(1,39) = 255.35, MSE = 6.61, p < .01). This shows that the short for condition lead to subtractive coercion.

The long for condition triggered the computation of iterative readings: 96.1% of the “yes” answers were “repeatedly”. In comparison, in the short in condition 95.1% of the “yes” answers were “only once”.

The semelfactives received 100% “yes, the event happened” responses in both conditions. Semelfactives modified by ago-adverbials had “only once” judgments
in 93.8% of the time, whereas semelfactives modified by for were in 94.6% of the time judged to express repeated events indicating the computation of iterative readings. ANOVAs revealed that the latter difference was significant \( F(1,23) = 2227.83, \text{MSE} = 9.36, p < .01; F(1,19) = 985.39, \text{MSE} = 7.83, p < .01 \).

To sum up, the pretest findings show that the materials have the intended meanings. In the accomplishments of Experiment 10a, short for-adverbials cause subtractive coercion. Long for-adverbials trigger iterative readings. The semelfactives also show a strong coercion effect: with ago-adverbials they are unequivocally interpreted as punctual events, but when modified by a for-adverbial they receive an iterative interpretation.

8.2 Iterative readings: Experiment 10a and b

Given the strong coercion effects showing up in the readings, do we find difficulty during online comprehension? Perhaps the materials used in Experiment 5 and 6 were too weak to elicit coercion effects. This is possible because they were not explicitly controlled for the interpretation subjects got. The materials in the present experiment should, however, be strong enough. The pretest showed that the materials have the intended meanings. Furthermore, nearly twice as many items per condition were included in the present experiment than in the previous ones.

8.2.1 Methods

The experiment was a self-paced reading study with the same method as Experiments 5 and 6. After reading a sentence participants had to provide a judgment whether the sentence was sensible. Experiment 10a tested accomplishments in the four conditions (6a–d). Experiment 10b tested semelfactives in the two conditions (7a–b). The 60 experimental sentences were combined with 90 fillers of which 40 were nonsensical. The overall ratio of sensible to nonsensical sentences in the experiment was 2:1. Four lists were constructed according to a latin square design.

Participants

Forty students from Tübingen University (all native German speakers, 33 female, mean age = 22.6, min. 19, max. 29 years) participated in the experiment. They were naive to the purpose of the study and hadn't participated in any of the previous experiments. Each subject was paid 5 Euros for participation. The participants were randomly assigned to lists (ten subjects per list). An experimental session took between 25 and 45 minutes.

8.2.2 Results

Sensicality judgments

The judgments of clearly sensible and nonsensical fillers shows that the participants read attentively. The sensible distractors were judged sensible 90.1% of the time while nonsensical distractors were judged sensible 20.2% of the time. All judgments are depicted in Figure 8.2.

![Figure 8.2. Mean sensicality judgments in Experiments 10a and b (+95% confidence intervals)](image)

**Accomplishments:** short for-adverbials were judged sensible 87.0% of the time and short in-adverbials were judged sensible in 82.3% of all cases. This difference didn't turn out to be significant in paired t-tests \( t(39) = 1.46, p = .15; t(39) = 1.54, p = .13 \). Long for-adverbials were judged sensible 69.3% of the time and were significantly below the short in-adverbials \( t(39) = 2.65, p < .05; t(39) = 3.35, p < .01 \). Long in-adverbials were judged sensible only 22.8% of the time and didn't differ from the nonsensical fillers \( t(39) = 0.58, p = .57 \). In ANOVAs with duration (two levels: short vs. long) and adverbial (two levels: for vs. in) as within factors, the difference between the long in condition and the other conditions was reflected by a significant interaction \( F(1,39) = 79.56, \text{MSE} = 1.74, p < .01; F(1,39) = 87.23, \text{MSE} = 1.74, p < .01 \) and by significant main effects both of duration \( F(1,39) = 157.26, \text{MSE} = 5.97, p < .01 \) and adverbial \( F(1,39) = 84.59, \text{MSE} = 2.63, p < .01 \).

**Semelfactives:** Semelfactives modified by for were judged sensible 83.0% of the time compared to 87.5% “yes” for semelfactives modified by ago. In paired t-tests, this difference wasn't significant \( t(39) = 1.45, p = .16; t(19) = 1.54, p = .14 \). Analyses of
answer times revealed no significant differences between conditions (all Fs < 1.5). The mean answer times varied between 950 and 1000 ms.

**Reading times**

**Accomplishments:** The reading times in the four accomplishment conditions are depicted in Figure 8.3. The adverbials were controlled for length, so in all subsequent analyses I report the raw reading times. These were corrected for outliers by trimming reading times that were below 200 ms or above 2500 ms.

At the adverbial, short *for*-adverbials were read fastest with a mean of 904 ms, short *in* had a mean Rt of 936 ms, long *for* 1000 ms and long *in* was read with a mean Rt of 1042 ms. ANOVAs revealed a significant main effect of duration ($F_1(1,39) = 18.35$, $MSE = 409910.64$, $p < .01$; $F_2(1,39) = 14.52$, $MSE = 409910.64$, $p < .01$) which is due to the fact that the long adverbials took longer to read than the short adverbials. The interaction between duration and adverbial was marginally significant in the subjects analysis, but wasn't reliable by items ($F_1(1,39) = 3.84$, $MSE = 53805.89$, $p = .06$; $F_2(1,39) = 2.13$, $MSE = 53805.89$, $p = .15$). Also, there was no reliable main effect of adverbial ($F_1/2 < 1$). Planned comparisons revealed that short *for*-adverbials didn't reliably differ from control ($t_1(39) = 1.25$, $p = .22$; $t_2(39) = 0.96$, $p = .34$), but that long *for*-adverbials took longer to read than control ($t_r(39) = 3.44$, $p < .01$; $t_2(39) = 2.92$, $p < .01$). Long *for*-adverbials didn't differ significantly from the implausible long *in*-adverbials ($t_1(39) = 1.73$, $p = .09$; $t_2(39) = 1.03$, $p = .31$).

**Semelfactives:** The reading times of coerced and non-coerced semelfactives are depicted in Figure 8.4. Paired t-tests revealed that the two conditions didn't differ in reading time at any segment (all $t_{1/2} < .5$).

> At the following segment, the implausible long *in* condition had slower reading times (564 ms) than the other conditions (short *for*: 485 ms, short *in*: 472 ms and long *for*: 509 ms). This difference was reflected by a significant interaction between duration and adverbial ($F_1(1,39) = 8.66$, $MSE = 46291.02$, $p < .01$; $F_2(1,39) = 9.16$, $MSE = 46291.02$, $p < .01$), a significant main effect of duration ($F_1(1,39) = 38.34$, $MSE = 133974.84$, $p < .01$; $F_2(1,39) = 36.67$, $MSE = 133974.84$, $p < .01$) and a marginal effect of adverbial ($F_1(1,39) = 3.36$, $MSE = 17365.97$, $p = .08$; $F_2(1,39) = 3.21$, $MSE = 17365.97$, $p = .08$). Paired t-tests revealed that long *for* was slower than short *in* ($t_1(39) = 2.58$, $p < .05$; $t_2(39) = 2.39$, $p < .05$) and long *in* was slower than short *in* ($t_1(39) = 5.56$, $p < .01$; $t_2(39) = 6.87$, $p < .01$). The numerical difference between short *for* and short *in* wasn't reliable ($t_1(39) = 1.17$, $p = .25$; $t_2(39) = 0.92$, $p = .36$). At later segments there were no significant differences between conditions (all Fs < 1).

8.2.3 Discussion

In the experiment, both subtractive coercion and iterated semelfactives were read as fast as non-coercing controls indicating that they were not difficult to process. The present study thus replicates the lack of coercion effects of Experiments 5 and 6.
Iterated accomplishments, however, slowed down reading pace just like the aspectually implausible long in condition. In contrast to the latter condition, the slowdown in iterated accomplishments is due to local aspecual mismatch that can eventually be repaired as indicated by the sensicality ratings. These effects demonstrate that the study was in principle sensitive to local and global difficulty in aspectual processing. This raises the question what makes subtractive coercion and abstract type shift so easy to process.

Subtractive coercion

In the discussion of Experiment 5 I have already speculated about the computational processes that underly subtractive coercion. Based on the present findings we can get a little bit more concrete. Let’s consider the subtractive coercion condition in (8) once more.

\[(8) \text{Der Bauarbeiter belud die Schubkarre eine Stunde lang, dann wurde er woanders eingeteilt.}\]

then was he somewhere else assigned

then he was assigned to another duty.

In the pretest, subtractive coercion sentences received less than 20% attainment of the goal judgments compared to nearly 80% in the control condition. This shows that readers interpreted sentences like (8) as not implying or implicating that the culminating event did happen. I will derive the meaning of (8) phrase-by-phrase using EC. I start with (9) which contains the integrity constraint (9a) and the definition in (9b) of the German perfective accomplishment 

\[
\text{den Schubkarren beladen.}
\]

To make integrity constraint (9a) true, the reader has to compute a discourse model where there were two past events. The first consists in a loading activity that was stopped after one hour. The second is an assigning event. Closed world reasoning ensures that the model is minimal, hence models are ruled out in which the activity was resumed after the stop event. This means that the culmination is supposed not to happen. This perfectly fits the judgment results.

In the controls instead of for an in-adverbial was used. Let’s compare the necessary steps required in subtractive coercion to those required in the control condition in order to determine their relative difficulty. In one hour makes reference to a finish event that occurred one hour after the preparation had started. Unifying it with \(\text{finish}_{\text{load}}\) yields an accomplishment with a preparation that lasted for one hour. Since there is no information about stop, it is assumed to be false. The derivation is analogous to subtractive coercion with the sole difference that the disjunctive formula among the defining conditions of the accomplishment is verified via \(\text{finish}\) instead of stop. The \(\text{then}\) clause is processed exactly like in the subtractive coercion example. It introduces an assigning event happening after the first event. The comparison shows that in terms of algorithmic complexity subtractive coercion is as easy as non-coercing controls. Its incremental interpretation, however, involves a non-monotonic update of the discourse model.

In languages other than German where perfective accomplishments have a more restricted meaning, the processing of subtractive coercion may be more difficult. In English, for example, the progressive is grammaticalized. Choosing between a progressive and a non-progressive form allows the speaker to disambiguate between accomplishments viewed as ongoing processes and accomplishments conceptualized as a complete whole. The latter seem to necessarily include all the parts that make up the accomplishment. If this is indeed the case, subtractive coercion of English perfective accomplishments modified by for-adverbials is predicted to be more difficult than subtractive coercion in German. In a self-paced reading study comparing subtractive coercion in German and English it was shown that this prediction is in fact borne out by reading time data (Bott 2009).

Iterative semelfactives

The pretest results show that the point action verbs were well chosen. In the control condition they were uniformly judged to denote punctual events, but when combined with a for-adverbial the meaning was shifted into an iterated process. Like in Experiment 6, coercion wasn’t reflected by enhanced processing load. The results are in line with the findings of Pickering et al. (2006) who also didn’t find difficulty in abstract type shift.

Does this mean that iterative readings come for free in general? I don’t think that this is the case. Comparing iterated semelfactives with iterated accomplishments shows
that there are differences in difficulty depending on aspectual class: the computation of iterative readings of accomplishments leads to considerable processing difficulty indicated both by higher rejection rates and by slower reading times.

How can the influence of aspectual class be explained? A semelfactive expresses a single event and doesn’t introduce any process fluent into the discourse model. As a consequence, the yet abstract process fluent introduced by the for-adverbial cannot be unified with any part of the existing representation. This triggers abstract type shift. Since type shifting is very frequent in natural language it is not difficult to perform and coercion goes unnoticed.

With accomplishments, the situation is different. In principle, accomplishments contain a process fluent that the long for-adverbial can be composed with. But doing so yields a implausible reading which has to be revised by the processor. The comprehender has to choose another route through the aspectual transition network of Moens & Steedman (1988). I will come back to iterative accomplishments in the next chapter because they are particularly interesting for evaluating different semantic theories on aspectual reinterpretation. For now, it suffices to see that there is reason to expect differences in difficulty between aspectual classes with respect to the computation if iterative readings.

To sum up, in the present experiment neither subtractive coercion nor abstract type shift caused difficulty in online processing. The results thus replicate the findings of the previous experiments (Exp. 5 and 6). Since “replication is the best statistic” (Luck 2005) the results provide strong evidence against both the Non-Monotonic Update Hypothesis (NMUH) and the Abstract Type-Shift Hypothesis (ATSH).

8.3 Summary and conclusions

In this chapter I have summarized the results of the experiments that have been reported in this monograph. Each was designed to test one or more hypotheses about aspectual processing and aspectual reanalysis. Taken together, the results can be used to make more specific the general model of aspectual reanalysis that was outlined in Chapter 3. The specified model is depicted in Figure 8.5.

Here is a brief summary of the model. Aspectual composition of a temporal modifier with an Aktionsart starts when a minimal verb-argument structure is complete. The Aktionsart is determined based on probabilistic information like lexical frequency of the verb and information from preceding discourse.

Composition can fail in one of three ways. First, the situation model may contain too little information for fully successful unification of modifier variables with constants contained in the situation model. If this is the case, additive coercion is required to enrich the model with additional eventualities. Second, the Aktionsart may be of

Figure 8.5. The specified model of aspectual processing

the wrong ontological type leading to a mismatch. If abstract-type shift is applicable the Aktionsart can be changed into the required type making composition possible. Finally, the Aktionsart may be more complex than needed. Subtractive coercion allows the comprehender to focus on parts of an eventuality by limiting it to a more restricted sense.

The top part of the model contains a plausibility check of the coerced meaning. If coercion yields an implausible meaning, the processor has to reanalyze the sentence once more by choosing another route. This is likely to be what we saw happen in the iterative accomplishments of Experiment 10a. Let’s step through an example. The processor has parsed an accomplishment (e.g. the worker loaded the wheelbarrow) which is modified by a long for-adverbial (e.g. for twenty years). Composition isn’t fully successful because the accomplishment is too complex to serve as semantic input for the adverbial. The comprehender thus has to follow the subtractive coercion route leading to an implausible interpretation of the sentence: a single loading activity that lasted for twenty years. This interpretation is at odds with world knowledge about typical events of loading a wheelbarrow. After checking the plausibility of the computed meaning, the sentence has to be reanalyzed again. This time another type
of coercion has to be chosen. The only remaining option is abstract type shift\(^1\) and the result is an iterative reading of the accomplishment. The slow reading times of iterated accomplishments suggest that failed reanalysis is costly. This means that the first reanalysis is free,\(^2\) but revising this meaning is expensive.

The specified model is a coarse-grained description of what the derivations of Chapter 3 spelled out formally in the combined DRT/EC framework. The model is a sketch of a computational theory in the sense of Marr (1982) while DRT/EC is at the more fine-grained level of representation and algorithm. In principle, DRT/EC allows us to formalize many different computations yielding the same meaning. Given the one-to-many relation between computational theory and representation/algorithm, we have to ask whether the results provide evidence for or against a particular computation, for instance the ones that have been proposed in the derivations of Chapter 3. More generally, what are the implications of the experimental findings with respect to the semantic coercion theories that have been presented in Chapter 2? Relating the results to theories and studies from the literature will be the topic of the next and final chapter.

---

1. Additive coercion is impossible because an accomplishment is already maximal.
2. This is only true for subtractive coercion and abstract type shift. Additive coercion is costly right from the start.

CHAPTER 9

Summary and conclusions

In this final chapter I will relate the experimental findings to semantic theories of aspectual reinterpretation and to the studies reported in the psycholinguistic literature on coercion. Although none of the semantic theories described in Chapter 2 was intended as a processing theory, we will see that psycholinguistic data can be used to evaluate the psychological validity of these accounts. At the same time, interesting predictions for psycholinguistic investigation can be derived from semantic theory. I will consider the findings of Experiment 10a again to demonstrate the mutual benefit of psycholinguistic investigation and formal semantic theorizing.

9.1 Relating the findings to formal semantic accounts of aspectual reinterpretation

In Chapter 2 I classified formal semantic theories on aspectual coercion into the following types: (1) operator-based accounts, (2) underspecification accounts and (3) planning accounts. Can the obtained results be used to decide between these theoretical alternatives? I will look at each of them in turn.

9.1.1 Operator-based accounts

Operator-based accounts assume that in aspectual coercion the semantic representation is repaired by plugging an appropriate type-shift operator into a semantic representation which would otherwise be ill-formed. The operator buffers away the type mismatch. Understood as a processing theory, an operator-based account predicts two things to happen in aspectual coercion. In trying to compose the coercing stimulus with the existing representation a type mismatch occurs. An appropriate type-shift operator is selected and fit into the semantic representation.

This two-step procedure is the same no matter what kind of coercion the comprehender is dealing with. In additive coercion, for instance, an operator is needed that adds an eventuality to the existing representation whereas in subtractive coercion the operator needs to input a complex eventuality and output a simpler situation type. Semantically, both types of coercion are semantically treated in the same way. As a
consequence, they might be expected to be similarly difficult to process. The differential effects obtained for additive coercion versus subtractive coercion and abstract type shift provide evidence against this prediction. Operator-based accounts have no explanation for this difference. An ad hoc explanation would be the following: Perhaps it is more difficult to come up with an “additive” operator than with a “subtractive” or an “iterative” one. This can’t be right either. Although iterative semelfactives didn’t cause difficulty, iterative accomplishments in Experiment 10a did. This is completely unexpected given the above explanation.

A second piece of evidence against operator-based accounts comes from the evoked potentials of Experiment 9. This study found no evidence for a semantic mismatch in aspectual coercion. The aspectual mismatch condition elicited a P600 effect but the coercion condition did not. There was no N400 effect indicating difficulty in semantic processing, either. Coercion only elicited a late working memory LAN. This suggests that additive coercion didn’t involve mismatch detection but only lead to a smooth update of the aspectual representation with a yet missing eventuality.

To sum up, operator-based accounts fail in two respects. First, they cannot explain differential processing difficulty in different kinds of coercion. Second, the neuropsychological correlates of additive coercion are more in line with a single step procedure than with a two-step model involving match detection and repair.

9.1.2 Underspecification accounts

Underspecification accounts have been proposed both in the semantic (Egg 2005; Dölling 2003a) and in the psycholinguistic literature (Pickering et al. 2006). Although sharing the same label, they mean quite different things. Semantic underspecification accounts assume that before computing a fully disambiguated meaning an underspecified representation is constructed which is the semantic representation proper. This underspecified representation is then disambiguated relying on pragmatic information. Semantic underspecification accounts are thus fully consistent with incremental word-by-word processing of specified representations. By contrast, the psycholinguistic underspecification account by Pickering et al. (2006) assumes that the processor sticks to the underspecified representation until disambiguating information is available which may be some words later in the sentence or discourse. Since the latter account is easier to test, I will look at it first.

Psycholinguistic underspecification

Pickering et al. (2006) assume that aspectual properties remain underspecified during ordinary comprehension. From their underspecification account it follows that after an aspectually ambiguous region disambiguation should be easy no matter what Aktionsart is required. This prediction was tested in Experiment 2. The results show that as soon as comprehenders had processed the subject and an aspectually biasing verb, they decided on an Aktionsart. This was indicated by the interaction of aspectual bias of the verb with subsequent adverbial modification by for- versus in-adverbs. The findings thus disconfirm aspectual underspecification in the sense of Pickering et al. (2006) and provide evidence for immediate commitment to aspectual class.

Semantic underspecification

Semantic underspecification accounts are very hard to evaluate because the underspecified representation is an intermediary construct that may immediately be specified. Moreover, the specification stage is hardly spelled out in these theories making it very difficult to make clear predictions about the actual computations. The findings of Experiment 2 are fully consistent with an incremental version of fully specifying the underspecified representation.

Semantic underspecification accounts are by definition two-step theories. With respect to the computation of the underspecified representation coerced and non-coerced sentences are in principle indistinguishable. When we compare the processing of coerced and non-coerced sentences only the pragmatic specification stage is relevant. Only during this step, coercion may become more difficult. If functor and argument are matching in type they can immediately be brought together by applying the identity function. By contrast, in coercion an operator is needed which intervenes between the two. Thus, specifying a coerced sentence involves adding an operator to the representation while non-coercing controls can do without. Selecting and integrating this operator could enhance working memory load. Semantic underspecification along the lines of Egg (2005) and Dölling (2003) thus, at first sight, fits rather nicely with the results of the ERP study (Exp. 9).

However, two findings are hard to explain under a semantic underspecification account. The first is differences in difficulty between the three types of coercion. In all of them some operator is required. Why should additive coercion be more difficult than the two other types of coercion? Just like operator-based accounts, semantic underspecification has no obvious explanation for this. The second finding that is problematic for underspecification are iterative accomplishments of Experiment 10a versus iterative semelfactives of Experiments 6 and 10b. If specifying the representation of a semelfactive modified by a for-adverbial can be easily obtained by plugging in an iteration operator, iteration of an accomplishment is expected to be just as easy since it involves basically the same operation. Why do we observe differences in processing difficulty among the aspectual classes? Semantic underspecification cannot explain these differences.
Planning accounts

I therefore favor the third kind of semantic accounts on aspectual coercion and here especially EC for the following reasons. First, a planning account like EC is cognitively grounded in that it provides the psycholinguist with an exact metric of difficulty (= number of applications of the fixed-point operator until the least fixed point is reached, see Appendix A) and with a flexible, algorithmic theory of meaning which can model a variety of alternative computations. Second, this type of theory is consistent with all the findings of the experiments that have been reported in this book. Finally, we can derive interesting predictions about the computational processes underlying different kinds of coercion. I have already demonstrated the first point in the derivations in Chapters 2 and 3 and in Appendix A. In the current section I will elaborate on the second point and in the following section I will turn to predictions that haven’t been tested yet.

Let’s briefly go through the reported experiments and see how the combined DRT/EC approach of Chapter 3 can deal with the results.

- **Experiment 2**: Using a combined DRT/EC framework offers the possibility to compute situation models incrementally by applying the DRS construction rules to every incoming word. Consider an experimental sentence of Experiment 2 like *John jogged in five minutes*. The model first gets updated with the fluent *John*, then the past activity *jog(x)* is added and *x* is unified with *John*. Integrating the in-adverbial leads to difficulty because it refers to a *finish*-event that isn’t contained in the model yet. Therefore, processing the adverbial leads to difficulty that is due to additive coercion (activity into accomplishment). The results of Experiment 2 can easily be accounted for by the DRT/EC framework.

- **Experiment 3**: The same holds for the findings of the context experiment which provided evidence that the processor makes immediate use of discourse context in determining the aspectual properties of a sentence. In the DRT/EC framework an in-adverbial can, for instance, simultaneously refer to both an activity fluent mentioned in the current sentence and a *finish*-event that was contained in preceding context. What is needed here is some kind of anaphora resolution and this is exactly what DRT was invented for. To see how it works I will briefly go through the four conditions tested in the experiment.

  - Telic contexts followed by *John jogged in thirty minutes*: The defining conditions of *in*(e, 30min.) consist of *Happens(finish, t)* and *HoldsAt(time, 30min., t)*. The former is satisfied by unifying *finish* with the path argument from the preceding discourse. The latter is satisfied by unifying *f* with the activity fluent *jog*. Processing of the target sentence is thus predicted to be as easy as combining the *in*-adverbial with the accomplishment. This prediction matches the results.

  - Telic contexts followed by *John jogged for thirty minutes*: For-adverbials are defined via an activity and a contextually given stop event. The former is contained in the target sentence. The latter can be unified with the path argument from the preceding discourse. The interpretation of the sentence is thus exactly the same as in the previous condition with an *in*-adverbial. However, the target sentence makes a weaker statement than would have been possible using *in*. Pragmatic reasoning à la Grice (1967) forces the reader to dissolve unification of *stop* with the path argument and to instantiate *stop* with some other event not contained in the discourse. Suspending the pragmatically unlicensed unification will enhance processing load. This is exactly what was found in the experiment.

  - Atelic contexts followed by *in vs. for*: These two conditions used neutral control contexts. The results match those of isolated sentences without context.

  - **Experiments 4a, 5 and 6**: The three experiments tested difficulty associated with different kinds of coercion. Only additive coercion caused difficulty. The derivations in Chapter 3 already suggested differential effects of the sort they were observed in the three experiments.

  - **Experiments 4b, 7 and 8**: The processing domain of situation aspect is rather large. Coercion and mismatch effects were only obtained when the coercing/mismatching stimulus was composed with a complete verb argument structure. In the slightly modified DRT framework I am using, Aktionsart is only specified at the sentential level. It depends on the values of three lexical features [±ADDTO], [±DUR] and [±SQA] which are brought into the representation by the verb and its arguments. The details have to be spelled out but for now it suffices that Aktionsart is only specified at the point when all pieces of relevant information

1. For instance, the analysis makes false predictions for intransitives, but only works for transitive verbs.
The ERP study served two goals: investigating the brain 
processing difficulty of experimental findings. The account performs better than any of its competitors. In the 

- Experiments 9a and b: The ERP study served two goals: investigating the brain 
signature of additive coercion and comparing the evoked potentials elicited by 

- It was shown that additive coercion is resolved by updating the situation 
model with an additional eventuality that has to be determined by extralinguistic means. Enriching the representation by coming up with a list of candidates and choosing the most plausible among them was predicted to enhance working memory load. Since the representation is smoothly updated, no mismatch detection was predicted to occur. We found exactly this pattern of results in the evoked potentials. While aspectual mismatch was indexed by a P600, additive coercion lead to a working memory LAN but not to a P600.

- The second prediction of the DRT/EC framework was about the processing domain of tense and Aktionsart. While Aktionsart involves a bigger domain, tense is computed on the basis of the tensed verb alone. Experiment 9b compared tense violations in adverbial-verb constructions with syntactically identical controls. Tense violations elicited a P600 effect immediately at the verb. This again confirms the prediction made by DRT/EC.

- Experiments 10a and b: The last experiment investigated processing difficulty of iterative readings in different Aktionsarten. It tested the processing of abstract type shift within semelfactives and accomplishments. While the former were as easy as non-coercing controls, the latter caused severe processing difficulty. The two derivations of abstract type shift in Chapter 3 are fully consistent with this pattern of results. In deriving iterated accomplishments it was hypothesized that first an implausible reading is computed which then has to be revised and reanalyzed. By contrast, in iterated semelfactives abstract type shift should be the only option and should be computed right from the start.

To sum up, the predictions of the DRT/EC framework were fully supported by the experimental findings. The account performs better than any of its competitors. In the next section, I will use DRT/EC to derive predictions about ERP effects that may be expected in the computation of iterative readings in accomplishments.

### 9.2 Predictions for psycholinguistic experiments

In an algorithmic semantic theory like DRT/EC meanings are derived step by step. I used this feature of the theory to predict the functional processes in additive coercion. Let’s try to account for the functional correlates of the observed disruption in iterated accomplishments. I will use the semantic derivation of iterative accomplishments in Chapter 3 to derive predictions about the brain responses in this construction. We will see that DRT/EC predicts another brain signature in this instance of coercion than in additive coercion.

The derivation involved reanalysis triggered by plausibility. Processing should thus be similar to that of the PP-attachment ambiguity in (11).

(11) The cop caught the thief with the **binocular**.

In this example, the PP *with the binocular* can be either attached to the VP yielding an instrument reading or to the NP. In English, VP attachment is preferred over NP attachment. When the processor encounters *with* it will initially attach the PP to the VP. But since a binocular is a bad instrument for catching, this reading turns out to be implausible. Instead, the more plausible NP attachment reading has to be chosen. To do so, both the syntactic and the semantic representation have to be reanalyzed.

The computation of abstract type shift in accomplishments was modeled exactly analogously to reanalysis in (11). The only difference is that iterative accomplishments only require semantic reanalysis. Given their similar derivations, the ERP effects in (12a) versus (12b) and (12c) versus (12d) are expected to be alike.

(12) a. The cop caught the thief with the **binocular**.
    b. The cop saw the thief with the **binocular**.
    c. For twenty years the mountaineer **climbed the mountain**.
    d. In twenty minutes the mountaineer **climbed the mountain**.

Based on what we know from the syntactic literature, we can speculate about what to expect in (12a) versus (12b). In an ERP study, Hare, et al. (1999) investigated sentences like (13a) and (13b).

(13) a. The priest believed the **doctrine** despite its problems **was** still fundamentally sound.
    b. The priest believed the **car** despite its problems **was** still fundamentally sound.
These garden-path sentences involve a direct object/complement clause ambiguity with the direct object interpretation being preferred over the complement clause interpretation. While the doctrine is a good direct object of believe, the car is not. NP type thus provides a plausibility bias towards the dispreferred reading in (13b). At the auxiliary was, both sentences are syntactically disambiguated towards the dispreferred interpretation. Hare et al. found two effects. Implausible direct objects showed an N400 effect relative to plausible direct objects. Further, at the disambiguating auxiliary of the verb phrase, both conditions elicited a P600 that continued through the end of the sentence but this effect was much larger in the plausible condition (doctrine...was). Thus, plausibility modulated the P600.

Given these results, in (12a) relative to (12b) a biphasic N400-P600 ERP waveform is expected, too. The N400 effect should follow from lexical access of an implausible instrument NP and the P600 from revision of the initial analysis. If the proposed analogy between syntactic and semantic reanalysis in iterative accomplishments holds, the difficulty in (12c) versus (12d) should also be reflected by a combined N400/P600 ERP waveform.

If this prediction turns out to be correct, the brain signatures of additive coercion and iterative accomplishments are fundamentally different. Additive coercion is indexed by a working memory LAN whereas iterative accomplishments should elicit a combined N400/P600 pattern. This offers the fascinating possibility to directly read off computational semantic processes from the evoked potentials. Using the DRT/EC framework, we can thus derive very explicit hypotheses about the cognitive processes in semantic interpretation.

9.3 Relating the findings to psycholinguistic studies on coercion

Turning to psycholinguistic work on coercion I will relate the findings to existing studies on aspectual and complement coercion. This will give us a more complete picture of coercion at the lexical level (complement coercion) and at the sentential level (aspectual coercion).

9.3.1 Studies on aspectual coercion

Previous research on aspectual coercion has reported conflicting evidence. Early studies found difficulty in abstract type shift using methods that involved a secondary task (Piñango et al. 1999; Todorova et al. 2000a; Todorova et al. 2000b; Piñango et al., 2006). However, these results have been questioned by Pickering et al. (2006) who failed to replicate them. Husband et al. (2006) and Brennan & Pylkkänen (2008a) report reading time results and MEG data that provide evidence for difficulty in abstract type shift.

However, the studies are hard to compare because the materials included sentences of various Aktionsarten. The only study that used normed materials which were clear instances of the class of semelfactives was Brennan & Pylkkänen (2008a). Unfortunately, as outlined in Chapter 2 many of the adverbials used in their control condition (after x time) made it necessary to update the temporal representation with a reference time not provided by the sentence making the results very hard to interpret.

The research presented in this monograph offers a systematic investigation of aspectual coercion. In line with Pickering et al. (2006), the typical examples of aspectual coercion (e.g., to jump for an hour) were found not difficult to process. Two other types of coercion, additive coercion and iterative accomplishments, did however enhance processing load. Thus, the findings of the dissertation provide evidence against under specification of Aktionsart. Instead they support graded difficulty of aspectual reanalysis depending on the computational processes which are required to compute the coerced meaning.

Two recent studies corroborate the claim that additive coercion causes difficulty. Brennan & Pylkkänen (2008b) presents reading times and MEG data on a subtype of additive coercion, namely coercing a state into an accomplishment. They compared coerced sentences like (14a) to controls such as (14b).

(14) a. Within a half hour, the kids distrusted the quiet librarian.
   b. For no good reason, the kids distrusted the quiet librarian.

In the coercion condition (14a), readers have to shift the stative meaning into an accomplishment. To do so, a preparatory process and a culmination have to be added to the stative situation. Reading times were longer in (14a) than in (14b). In MEG, coercion elicited a peak in anterior midline activity about 400ms after the onset of the verb. This is the same MEG component that was reported in an earlier study in connection with complement coercion (Pylkkänen & McElree 2007) and has also been associated with social cognition. Since this is a different construction type involving additive coercion, their findings corroborate and generalize the claim that additive coercion is costly. Moreover, inferring the right kind of preparatory process can plausibly be argued to elicit enhanced activation in the AMF because social knowledge is needed to infer what kind of actions librarians may perform that cause kids to distrust them.

The second study reports reading time data on aspectual coercion in Japanese (Shengyan, et al. 2008). In their experiments, in- and for-adverbials were used to trigger additive coercion of activities and subtractive coercion of accomplishments. The results show increased reading times of adverbials triggering additive coercion. Thus, their findings support the results of this thesis. With respect to subtractive coercion, however, the findings seem to be incompatible. Reading times slowed down considerably when a for-adverbial was combined with a perfective accomplishment. How can the apparently conflicting crosslinguistic data be reconciled? Here is an attempt of an
explanation. Japanese has grammaticalized aspectual markers indicating imperfective and perfective viewpoint aspect. Perfective accomplishments are unambiguously telic and thus entail the culminating event. It thus may be the case that in Japanese only the second, more costly derivation of subtractive coercion is possible that was outlined in Chapter 3. This makes reinterpretation much harder than in German. To test this explanation, it would be highly instructive to study subtractive coercion in a language like English which also has a grammaticalized distinction between imperfective and perfective forms. If the above explanation is correct, subtractive coercion of a perfective accomplishment into an activity is predicted to be difficult in English, too. In a cross-linguistic reading time study this was actually the case (Bott 2009).

9.3.2 Studies on complement coercion

The findings of the present experiments can best be related to studies on complement coercion by comparing ERP effects elicited by nominal and aspectual coercion. A recent ERP study by Kuperberg et al. (2009) used an experimental design maximally similar to that of Experiment 9a. They investigated the ERP signatures of sentences (15a–c).

(15) a. The chef finished the chicken before the main course.
   b. The chef cooked the chicken before the main course.
   c. The chef astonished the chicken before the main course.

(15a) involves complement coercion, (15b) is a non-coercing control and (15c) is a mismatch condition. Kuperberg et al. found two effects. At the complement noun, coerced and mismatching nouns evoked a N400 effect of similar magnitude relative to control. This is consistent with the findings of Choma (2006) who also reports a N400 using basically the same design. At the sentence-final word, complement coercion produced a prolonged anterior positivity relative to mismatch and control.

In Experiment 9a, aspectual processing evoked a fundamentally different brain response. Let’s consider the differences and their implications for theories of both coercion types.

- Complement coercion elicits a N400. Taking the N400 as an index of lexical access this difference can be interpreted to support lexical type shift in complement coercion. In this type of coercion, the lexical meaning of the complement noun has to be transformed to meet the input requirements of the verb. The N400 data fit nicely with the account of complement coercion by Pustejovsky (1995) who models repair within the lexicon. However, aspectual coercion and aspectual mismatch do not elicit a N400. This kind of coercion isn’t resolved at the lexical level but affects the whole sentence or even discourse context. Complement coercion and aspectual coercion should thus be treated differently.

Despite these differences, there are also similarities between complement coercion and aspectual coercion. The MEG studies on complement coercion by Pylkkänen & McElree (2007) and on additive coercion by Brennan & Pylkkänen (2008b) both report increased amplitudes in the anterior midline field (AMF). It has been shown that patients with lesions in the ventro-medial prefrontal cortex – the generator of the AMF – have difficulty with false belief tasks commonly employed in research on theory of mind (Rowe et al. 2001). Thus it seems that in both kinds of coercion inferring the appropriate kind of action depends on reasoning about the intentional stances of the agent with regard to the goal. For instance, in the classical example of complement coercion, the author began the book, the comprehender has to reason about the intentions authors typically have towards books (e.g. they plan to write them). Similarly, in Brennan & Pylkkänen (2008b)’s inchoative coercion and in additive coercion in general, the comprehender has to reason about what actions the agent may consider to reach the goal state. If this explanation is correct, the AMF doesn’t reflect language specific semantic composition but can be interpreted in a more general way as reflecting reasoning about the intentions and goals of the event participants.

To sum up, complement coercion involves reinterpretation at the lexical level. I suggest that the coerced meaning of the complement noun is computed using social cognition. In this type of coercion, lexical processing seems to interact with social cognition and/or theory of mind.

9.4 Relating the findings to models of sentence processing

So far, I have been discussing the findings only with respect to semantic and psycholinguistic accounts of aspect and coercion. In this section, I will relate them to models of sentence processing in general. Since there is a vast literature on the cognitive architecture employed in syntactic processing I will simplify a great deal and only focus on the extreme positions, syntax first accounts as compared to constraint satisfaction models.

For a long time, one big question in the research on sentence processing was how initial parsing decisions are made (for an overview see Pickering & van Gompel (2006)). Does the processor only use syntactic information or does it build the initial parse relying on all kinds of knowledge sources simultaneously? While syntax first
accounts assume the first alternative to be correct, constraint based models act on the assumption that all sources of information are taken into account immediately. In the last more than twenty years plenty of evidence has been obtained that besides syntax also non-syntactic factors (e.g. lexical frequency, discourse context, …) play an important role in determining the initial parse. However, these factors often do not entirely override basic preferences for particular types of structures showing that structural preference does play a key role during sentence processing.

The DRT/EC framework implicitly assumes modularity because DRS resolution starts with the syntactic tree. This means in modeling the results we have all the time been working with a syntax-first theory of semantic processing. The reader may ask whether the modular architecture of the theory can be psychologically valid given the solid evidence for constraint-satisfaction from the literature on syntactic parsing. Although this is a valid objection it does not dismiss EC as a processing theory. The syntax semantics interface could be modeled in fundamentally different ways. To illustrate this point, in the next section I will briefly mention a theoretical alternative.

9.4.1 Towards an immediacy model of semantic processing

In generative grammar, syntax is the only generating device. The compositional semantic interpretation is assigned on the basis of the syntactic analysis. This implies that any semantic processing theory couched in the framework of generative grammar is necessarily syntax-first. An implication of this type of theory is that semantic information cannot be processed in the absence of syntactic information – an assumption that has been challenged by psycholinguistic experimentation (Hagoort & van Berkum 2007).

An alternative, constraint-based architecture has been proposed by Jackendoff (1996) in which phonology, syntax and semantics are all generating devices which work in parallel and are linked via interface rules. Baggio et al. (unpublished manuscript) sketch a processing model based on Jackendoff’s proposal. Syntax and semantics are lexicalized with lexical items carrying both a syntactic and a semantic frame. Syntactic processing involves unifying nodes in different syntactic frames; semantic processing involves unifying variables in different semantic frames. The syntactic theory has been spelled out within the framework of tree adjoining grammar by Vosse & Kempen (2000) and has been related to brain imaging studies by Hagoort (2003a). The semantic component is based upon EC. Let’s look at the analysis of complement coercion in (16).

(16) The novelist began the book.

Baggio et al. (unpublished manuscript) assume that lexical items come with pieces of hierarchical structure having unification sites, that is a root node and foot nodes. In parallel, they enter their lexical meaning (an EC representation) into the semantic workspace with variables serving as semantic unification sites. (17a) is the syntactic frame and (17b) is the semantic frame of novelist2 containing the information that novelists are people who write books.

(17) a. NP
   det head modifier
   DP NP PP
   novelist

b. $\exists x. (\text{HoldsAt}(\text{book}(x), t) \land \text{Happens}(\text{write}(\text{agent}(y), \text{theme}(x)), t)) \rightarrow \text{HoldsAt}(\text{novelist}(y), t)$

The next word is began. It is associated with the syntactic frame (18a)3 and the semantic representation (18b).

(18) a. S
   subject head dir. object modifier
   NP VP NP PP
   begin

b. $\text{Initiates}(\text{begin}, \text{activity}(\text{agent}), t)$

The frames of novelist and begin can now be assembled. Syntactic unification will attach the root node of author to the subject NP foot node of begin combining the two trees into a single one. The semantic frames are combined by unifying agent with the novelist.

2. For simplicity, I leave out the representations of the and tense.
3. The example contains a local syntactic ambiguity. Begin can either be used with a nominal complement or with an infinitival complement. However, at the word following the verb it becomes already clear that begin is used with a nominal complement here. In complement coercion the critical word is the complement noun. At this word the sentence has already been disambiguated. Thus, for simplicity, I assume the syntactic analysis in (18a) to be the only one.
When encountering the next phrase, the root node of the book in (19a) will be attached to the direct object foot node of begin. In the semantic component, enriched composition is required because the lexical entry of book in (19b) is of the wrong type to be unified with activity.

![Diagram of the book node](image)

(19) a. NP
   det head modifier
   DP NP PP
   book

b. \( \text{HoldsAt}(\text{pages}(x), t) \land \ldots \land \text{HoldsAt}(\text{cover}(x), t) \rightarrow \text{HoldsAt}(\text{book}(x), t) \)

Semantic enrichment proceeds as follows: the accomplishment write(agent, theme) is activated by the lexical representation of novelist. It can be unified with activity. Write(agent, theme) introduces the following two conditions (among the other defining conditions of the accomplishment scenario).

(20) \( \text{Releases}(\text{start}, \text{theme}(x), t) \rightarrow \text{Trajectory}(\text{write}(\text{agent}), t, \text{theme}(x + g(d)), d) \)

Unifying agent with the novelist and theme with the book yields the coerced interpretation. Thus, this semantic derivation of complement coercion is completely independent of the syntactic parse. Baggio et al. (unpublished manuscript)’s theory is a one stage model of semantic and syntactic parsing with both components generating representations in parallel.

The reader may thus wonder why I used the DRT/EC framework instead of the immediacy model. Although I’m fully sympathetic with Baggio et al’s model there are some reasons for this choice.

The first is that DRT is a well established semantic theory able to deal with a vast amount of semantic phenomena, for instance, quantifier scope or restrictions on anaphora resolution.

Second, the number of possible unification candidates is much more constrained in the DRT/EC framework. This is because syntax imposes locality constraints on the semantics. Consider, for instance, a simple sentence like the girl kissed the shy boy. What is expected to happen when the processor encounters the adjective shy? Supposedly the adjective will have a frame \( \text{HoldsAt}(\text{shy}(x), t) \) with \( x \) being a variable that has to be unified with some discourse referent. At that point, there is only one, namely the girl. Hence the semantic processor will unify the two yielding a semantic representation that is in conflict with the syntactic parse. The example shows that without syntactic guidance the semantic parser will (initially) consider far too many unification options. I think this is highly implausible. What seems to be more correct is that the syntactic information determines what may count as a local domain for semantic processing.

Finally, the authors themselves note that the model might even be ‘too incremental’. Constraining the semantic component, however, is far from trivial. By contrast, in DRT/EC delayed processing can be easily modeled by using features that are only specified at a particular phrasal level. It is an open question how incremental semantic processing really is.

This illustrates that EC is flexible enough to model both, one-stage and two-stage theories of semantic processing. It is the architecture of the syntax-semantics interface (interface rules versus DRT) that mediates the relation between syntax and semantics. Although it is possible to develop an immediacy model incorporating EC to work in parallel with the syntactic parser, I think there are still reasons for preferring EC in the the DRT/EC system which is a syntax-first theory of semantic processing.

### 9.5 Open questions

In this section I list a number of questions that have been raised along the way. I will discuss them in turn. First, I will list questions in connection with the experimental findings of the monograph. This will set an agenda for further research. I will close with a discussion of open problems in modeling aspectual processing.

#### 9.5.1 Questions for further research

In the introduction I emphasized the universal nature of the aspectual classes and implicitly made the assumption that aspectual processing is similar across languages. Subtractive coercion casts doubt whether this presumption is correct. Comparing German data with English (Bott 2009) and Japanese data (Shengyan et al. 2008) revealed cross-linguistic differences. This lead to the hypothesis that subtractive coercion can be difficult but difficulty depends on how grammatical aspect is realized in a given language. More generally, this line of reasoning leads to the question on how the aspectual system of a language does influence aspectual processing. For instance, is Aktionsart processed on a word by word basis in a language with a richer aspect system than German and grammatical means to express Aktionsart?

Another open issue concerns the localization of coercion effects in the brain. Planning accounts like EC predict that brain areas which are usually activated in planning tasks should be active in computing a situation model, too. In Chapter 7 I have speculated about brain areas that may be active in additive coercion deriving predictions from Hagoort (2005)’s neurolinguistic theory. Based on Hagoort’s model, a double
dissociation of aspectual mismatch and additive coercion is expected in processing the materials of Experiment 9a. While mismatch should lead to enhanced activity of the semantic unification space, additive coercion should lead to enhanced activity in control regions and in memory.

Finally, the online composition of events is a major desideratum. How do various pieces of linguistic information interact when an event representation is computed online? What is going on when the verb receives an argument but other arguments are still missing? Is the processor just waiting for the rest to come? Or should we think of this as making a situation more concrete starting with a rather general model which is narrowed down more and more until the comprehender eventually reaches a specified event representation? Although psycholinguistics has started to work on these questions they are still far from being answered.

9.5.2 Issues in modeling aspectual processing

Throughout this monograph I have tried to integrate the experimental findings into a computational model of aspectual semantic processing. I ended up with a processing model that allows us to derive fairly specific predictions. There are, however, some aspects that require further work.

First, a more systematic investigation of aspectual coercion is needed. So far, only a few routes in the transition network of Moens & Steedman (1988, see Figure 2.2) have been investigated. Take for instance an achievement like the retiree found his keys for twenty years. The sentence is semantically ill-formed although the specified model of aspectual reanalysis predicts that iteration should be possible. Intuitively, it is clear why the sentence is bad. Finding the keys will lead to a state of having them. This leads to a contradiction with the precondition of find, namely being in a state of not having them. However, this line of reasoning has to be spelled out formally.

Second, subtractive coercion has to be modeled by means other than a disjunction in the body of the defining statement of a perfective accomplishment. We need a more general device to handle aspectual semantic ambiguity instead of “hand-coding” it for every German perfective accomplishment.

Third, it is an open issue whether a one-stage or a two-stage theory of semantic processing should be adopted. Incrementality of semantic processing has to be tested using a wider range of clearly supralexical semantic phenomena. For example, the processing of scope ambiguities is of great interest here. Do we find evidence for immediate scope conflict in scopally ambiguous sentences? When is world knowledge taken into consideration?

This leads us to the last and most tricky question. In this book, EC was used to model the world knowledge of a comprehender. These “background assumptions” together with the linguistic input went into the computation of a situation model. But so far, I only used toy examples in which the convenient formula was selected beforehand. What happens if the knowledge base is scaled up to model real world knowledge? Probably, there will be many alternatives and often conflicting solutions. How does the comprehender select from them? Are all taken into consideration simultaneously or is it only the most salient that are considered? Further, how can salience be determined? To date, there is no answer to these questions.
References


Baggio, G., van Lambalgen, M. & Hagoort, P. Unification is the key to discourse processing. Unpublished Ms.


The Processing of Events


Poster presented at the 21st CUNY conference.


APPENDIX A

The semantics of logic programming and its implementation by recurrent neural nets

This appendix serves two purposes. First, it is intended as a brief introduction to the semantics of logic programming. For reasons of simplicity, the computational machinery is illustrated for propositional logic programming. Further, it will be demonstrated how computations of logic programming can be simulated using recurrent neural nets. They will be used to compute the minimal models of propositional logic programs. Second, the appendix is intended to demonstrate that (subtractive) coercion can be computed in different ways which can be distinguished in terms of their computational complexity. For the first computation, an accomplishment (which in propositional logic programming can be expressed by a clause of the form \( p_{\text{activity}} \land \neg \text{abnormal} \rightarrow q_{\text{culmination}} \)) is coerced into an activity \( p_{\text{activity}} \) right from the start. By contrast, an alternative computation is shown to be more complex and to involve a non-monotonic update of the minimal model. The comparison will further show how difficulty can be approximated in terms of algorithmic complexity.

The next section will provide an introduction to the semantics of logic programming and the following section will describe how model construction can be implemented by spreading activation networks.

A.1 Propositional logic programming

The logical form of a sentence is stated in logic programming, a fragment of propositional (or predicate) logic including negation as failure with a special non-classical semantics. The following definitions are from Stenning & van Lambalgen (2005, p. 936 ff.).

**Definition 6** A (definite) clause is a formula of the form \( \neg p_1 \land \ldots \land \neg p_n \rightarrow q \), where the \( p_i \) are either propositional variables, \( \top \) or \( \bot \), and \( q \) is a propositional variable. Facts

---

1. \( \top \) is used for an arbitrary tautology, and \( \bot \) for an arbitrary contradiction.
The Processing of Events

are clauses of the form \( T \rightarrow q \), which are abbreviated to \( q \). A definite logic program is a finite conjunction of definite clauses.

The closed world assumption is realized as “negation as failure”: \( \neg \phi \) is true if the attempt to derive \( \phi \) from the program \( P \) fails. The semantics for definite programs is stated in terms of three-valued logic, which has the truth values true (1), undecided (\( u \)) and false (0). Undecided means that the truth value can evolve either toward true or false. The truth tables for the connectives \( \neg, \land \) and \( \lor \) are provided in Figure A.1 (from Kleene (1951, p. 334)).

\[
\begin{array}{c|c|c|c}
  p & q & p \land q & p \lor q \\
  \hline
  1 & 1 & 1 & 1 \\
  0 & 0 & 0 & 0 \\
  1 & 0 & 0 & 1 \\
  0 & 1 & 1 & 1 \\
  u & u & u & u \\
  u & 0 & 0 & u \\
  0 & u & u & 1 \\
  0 & 0 & u & 0 \\
  \end{array}
\]

\[
\begin{array}{c|c|c|c}
  p & \neg p & & \\
  \hline
  1 & 0 & & \\
  0 & u & & \\
  u & 1 & & \\
  u & 0 & & \\
  \end{array}
\]

\[
\begin{array}{c|c|c|c}
  & p & q & p \land q \\
  \hline
  p & 1 & 1 & 1 \\
  \neg p & 0 & 0 & 0 \\
  \end{array}
\]

Figure A.1. Truth tables for three-valued connectives

The semantics of the nonmonotonic consequence relation \( \models \) is non-classical. If \( P \models \phi \) then \( \phi \) follows from \( P \) by negation as failure or by closed world reasoning. Instead of validity in all models only the preferred models are considered. The models of interest are captured by means of the following construction.

**Definition 7**

a. The completion of a program \( P \) is given by the following procedure:

1. take all clauses \( \phi_i \rightarrow q \) whose head is \( q \) and form the expression \( \lor \phi_i \rightarrow q \)
2. replace the \( \rightarrow \)'s by \( \leftrightarrow \)'s (the biconditional of classical logic)
3. this gives the completion: \( \text{comp}(P) \)

b. The nonmonotonic consequence relation \( \models \) is defined by

\[ P \models \phi \iff \text{comp}(P) \models \phi. \]

The completion of program clauses allows to reason from the premises to the conclusion. Consider the logic program \( P \) consisting of the clauses \( \{ p; r \rightarrow q; \perp \rightarrow ab \} \). Closed world reasoning as formalized in the completion yields \( \{ p; r \rightarrow q; \perp \leftrightarrow ab \} \) which is equivalent to \( \{ p; r \rightarrow q \} \) from which \( q \) follows. What happens when a disabling condition \( \{ r; r \rightarrow ab \} \) is added? The completion now becomes

\[ \{ p; r \land \neg ab \rightarrow q; \perp \lor r \leftrightarrow ab \} \]

which reduces to \( \{ p; r; ab; \neg q \} \). Here we see nonmonotonicity at work: the minimal model for a logic program with additional information may cancel an inference (namely \( q \)) that was true before. This kind of closed world reasoning plays the key role in interpreting imperfective accomplishments.

Minimal models can be effectively constructed. They are given by the last fixed points of a monotone operator \( T_P \) which rests upon a partial ordering \( \leq \) on truth values (\( u \leq 1 \) and \( u \leq 0 \)). By iterative applications of \( T_P \), a minimal model can be constructed in finitely many steps starting with the empty model (that is the model in which the truth values of all proposition letters are yet undecided). The neural nets in the next section mimic the behavior of the \( T_P \) fixed point operator. Minimal models will then correspond to the stable states of recurrent nets.

**Definition 8** A three-valued model is an assignment of the truth values \( u, 0, 1 \) to the set of propositions letters. If \( M, N \) are models, the relation \( M \leq N \) means that the truth value of a proposition letter \( p \) in \( M \) is less than or equal to the truth value of \( p \) in \( N \) in the canonical ordering on \( u, 0, 1 \).

Definition 8 defines an ordering on models which ranks models with a 0 or 1 truth value higher than models which have more \( u \) values. Constructing a model then has the intuitive meaning of further specifying a given model step by step. A fixed point is reached if the model is maximally specific.

**Definition 9** Let \( P \) be a program

a. The operator \( T_P \) applied to formulas constructed using only \( \neg, \land \) and \( \lor \) is determined by the above truth tables.

b. Given a three-valued model \( M \), \( T_P(M) \) is the model determined by

i. \( T_P(M)(q) = 1 \) iff there is a clause \( \phi \rightarrow q \) such that \( M \models \phi \)
ii. \( T_P(M)(q) = 0 \) iff there is a clause \( \phi \rightarrow q \) in \( P \) and for all such clauses, \( M \models \neg \phi \)
iii. \( T_P(M)(q) = u \) otherwise

**Definition 10** A fixed point of \( T_P \) is a model \( M \) such that \( T_P(M) = M \)

Once a proposition letter has been assigned 0 or 1 by \( T_P \), it retains that value throughout the model construction. If it has been assigned value \( u \) that value may change at a later stage. In their book van Lambalgen & Stenning (2008, p. 167) list the following two lemmas which are responsible for the efficient implementability in neural nets.

**Lemma 11** If \( P \) is a definite logic program, \( T_P \) is monotone in the sense that \( M \leq N \) implies \( T_P(M) \leq T_P(N) \)
Lemma 12 Let \( P \) be a definite logic program

1. The operator \( T_p \) has a least fixed point, obtained by starting from the model \( M_0 \) in which all proposition letters have the value \( u \). The least fixed point of \( T_p \) is called the minimal model of \( P \).
2. The minimal model of \( P \) is reached in finitely many steps \( (n+1 \text{ if the program contains } n \text{ clauses}) \).
3. All models \( M \) of \( \text{comp}(P) \) are fixed points of \( T_p \) and every fixed point is a model.

The following examples may help to clarify these notions. The first example is a logic program which contains a simple proposition \( [p] \) like it is needed to encode an activity.

The second example illustrates the construction of a model for an imperfective accomplishment \( [p; p \land \neg ab \rightarrow q] \). The last example illustrates the semantic effect of adding a disabling condition. In this example a disabling condition is added to an imperfective accomplishment \( [p; r; p \land \neg ab \rightarrow q; r \rightarrow ab] \).

- Let the definite program consist of \( [T \rightarrow p] \). Then \( M_1(p) = T_p(M_0)(p) = 1 \). The model \( M_1 \) is the minimal model.\(^2\)
- The second program is \( [T \rightarrow p; p \land \neg ab \rightarrow q; \bot \rightarrow ab] \). Then \( M_1(p) = T_p(M_0)(p) = 1 \). In the next iteration in addition to \( p = 1 \) and \( ab = 0 \) we get \( M_2(q) = T_p(M_1)(q) = 1 \). \( M_1 \) is the least fixed point.
- The last logic program I will consider is \( [T \rightarrow p; T \rightarrow r; p \land \neg ab \rightarrow q; r \rightarrow ab] \). The models contain the proposition letters \( p, q, r \) and \( ab \). In \( M_1 \) they receive the following values: \( p = 1, q = u, r = 1, ab = u \). The next iteration yields \( M_2(p) = 1, q = u, r = 1, ab = 1 \), hence \( \neg ab = 0 \). \( M_2 \) is the least fixed point: \( p = 1, q = 0, r = 1, ab = 1 \).

The number of iterations which are needed to arrive at the minimal model can be used to index complexity of the semantic derivation. Thus, the second example is more complex (= needs more computation) than the first and the third is even more complex than the second. An additionally nice feature of the described algorithm is that the construction of a minimal model can be modeled using neural nets.

A.2 The construction of minimal models using neural nets

The algorithm specified will be neurally implemented using recurrent neural nets as demonstrated in van Lambalgen & Stenning (2008, Chapter 8). Here are relevant definitions.

\(^2\) This model contains \( p \) (which has truth value 1) and nothing else.

Definition 13 A computational unit is a function with the following input-output behavior

1. Inputs are delivered to the unit via links, which have weights \( w \in [-1, 1] \)
2. The links \( x_i, \ldots, x_n \) can be excitatory or inhibitory: If an excitatory input fires the unit receives a positive amount of activation (+1), if an inhibitory input fires the unit receives negative activation (−1)
3. The quantity \( \sum_{i=1}^{n} x_i w_i \) is computed. If this quantity is greater or equal to a threshold \( \theta \), the unit outputs 1, if not it outputs 0.

Figure A.2. Representation of the node corresponding to a proposition \( p \)

The artificial computational units are also called McCulloch-Pitts-neurons. They transform an input into a binary output. McCulloch-Pitts-neurons are either active or inactive. Similar to their real counterparts, the artificial neurons are active at a certain point in time if they have received enough excitatory input one moment before. For ease of modelling, the time continuum is sliced into discrete time steps.

The artificial neurons may vary with regard to the number of excitatory inputs they need to become active. How much excitation is needed depends on the threshold \( \theta \). For instance, in case \( \theta \) equals 1 one excitatory input is enough to activate the neuron. Interestingly, such a neuron can be used to model the logical connective or. The neuron representing \( x_1 \lor \ldots \lor x_n \) will become active (that is true) iff at least one of the excitatory \( x_i \) is active (true). For the calculation of an “either or” connective, the solution cannot be achieved by a single neuron. It is necessary to introduce intermediate processing steps and to arrange neurons in layers through which activity flows serially. A neuron arrangement with an input layer and two subsequent processing steps can realize any event that can be described by a logical formula including the sentence connectives and, or and not (Kleene 1956).

Because I am using three-valued logic, I will extend the notion of a computational unit to the notion of a computational node. The underlying idea is that the values \( u, 0 \) and \( 1 \) can be expressed by pairs \((0,0),(0,1)\) and \((1,0)\) of which the first component stands for positive evidence and the second component represents negative evidence of a given proposition. Accordingly, each node consists of two neurons, namely a positive and a negative unit. The value \((1,1)\) is ruled out via coupled inhibitory links.
The following definition introduces the theoretical term of a node. Figure A.2 depicts a node.

**Definition 14** A node consists of two coupled computing units $U_i$ and $U_{-i}$. $U_i$ projects an inhibitory link to $U_{-i}$ and vice versa. A node is denoted by $(U_i, U_{-i})$.

How can three-valued negation and conjunction be realized in a network? Let’s consider negation first. Three cases have to be considered: The input node can have the values (0,0), (0,1) or (1,0). The respective outputs are (0,0), (1,0) and (0,1). Figure A.3 shows how three valued negation can be modeled through two intervening inhibitory units. If the input is (0,0) the inhibitory units will output (1,1) and the subsequent not-node will turn this into (0,0) again. If the input is (1,0), the inhibitory units will turn positive and negative values into (0,1). Similarly, if the input is (0,1).

![Figure A.3. Three valued negation](image)

Three valued and is realized by connecting n conjunct nodes to an AND-node via excitatory links. The threshold of the AND$_i$ unit is set to n while the threshold of the AND$_{-i}$ unit is set to 1. As one may check, this configuration yields the truth-table of three-valued and. Three valued or is constructed accordingly but this time the thresholds are different. While the OR$_i$ unit receives a threshold of 1, the OR$_{-i}$ unit has a threshold of n.

Now, everything is ready to define coupled recurrent networks which will be used to compute minimal models.

**Definition 15** A recurrent neural network of the relevant type is a set of nodes $U_i$ ($1 \leq i \leq n$) connected by links which may be inhibitory or excitatory. The network is arranged in at least three layers: an input, one or more intermediate layers and an output layer. And, or and not nodes can only appear in the intermediate layer. Each propositional letter appears once in the input and in the output layer. Further there is an excitatory link connecting the proposition letter in the input layer with its counterpart in the output layer. In addition to the units $U_i$ there are two bias units 1 and 0 which may be connected to some of the $U_i$ in the output layer. The bias units serve as input and encode what is known to be true and false in a given program P. They fire continuously.

Appendix A. Logic programming and neural nets

Let time range over positive integers. At each instant $t$, the global state $U_i(t)$ is computed. At time 0, all units except bias units are inactive (= have value u). To compute $U_i(t + 1)$ for each unit $U_i$, the activation $A_i(t+1)$ is computed with $A_i(t+1) = \sum_j x_i(t) w_{ij}$, where $j$ are the inputs to neuron $i$. $U_i(t + 1) = 1$ if $A_i(t + 1) \geq \theta_i$ and 0 otherwise.

**Definition 16** A coupled recurrent network consists of two recurrent networks such that for each unit in one layer there is a unique unit in the other layer with which it forms a node. The 1 bias node has only connections to the + sheet and the 0 bias node has only connections to the – sheet.

Let’s go through the examples of the last section by setting up coupled recurrent networks for the programs. To make the examples better comparable I will always introduce the same set of proposition letters {p, q, r, ab} into the neural nets. This has the side effect that it allows to model working memory as the activated part of declarative memory (like it has been proposed e.g. by Cowan (2005)). Thus, although all propositions are part of declarative memory, only some receive activation during interpretation.

The easiest case is the logic program which just consists of an activity [T → p]. Figure A.4 shows the coupled recurrent net into which this program translates.

![Figure A.4. Recurrent net representing the program {p}](image)

For simplicity, each node is depicted as a single graphical unit, but these graphical units have truth values u, 1 or 0. At time 0 all proposition letters have value u. At time 1, the positive bias node activates $P_{\text{output}}$ in the output layer. At time 2, activation spreads to the node $P_{\text{input}}$ in the input layer while $P_{\text{output}}$ still receives activation from the bias unit. This is already a stable state since at time 3 the net exhibits the same activation pattern. The minimal model can be read off the activation states of the propositional nodes and consists of p being true, while q, r and ab are undecided.

The next example is the logic program of an imperfective accomplishment {T → p; ⊥ → ab; p ∧ ¬ab → q}. The corresponding net is depicted in Figure A.5.
Again, at time 0 we start with the empty model, hence all propositions have value \( u \). At time 1, \( p_{\text{output}} \) is activated by the positive bias unit, while \( ab_{\text{output}} \) is activated by the negative bias unit. As a result, \( p_{\text{output}} \) is true and \( ab_{\text{output}} \) is false. In the next step, \( p_{\text{input}} \) and \( ab_{\text{input}} \) are activated. Because \( ab_{\text{input}} \) is false, \( \neg ab \) changes to true in the third step. In the fourth step, the \( \text{AND} \) node becomes true. Finally, steps five and six activate \( q_{\text{output}} \) and \( q_{\text{input}} \) respectively and the network reaches its stable state. The minimal model consists of \( p, q = 1 \), \( ab = 0 \) and \( r = u \).

The last example consists of an imperfective accomplishment with a disabling condition \( \bot \rightarrow p; \bot \rightarrow r; p \land ab \rightarrow q; r \rightarrow \neg ab \). The according network is shown in Figure A.6. The following list shows the activation flow through the network. After time step 8, the stable state is reached. The minimal model consists of \( p, r, ab = 1 \) and \( q = 0 \). The inference of the goal state \( q \) is no longer valid.

The examples involve an increasing number of time steps until the stable state is reached. The number of time steps needed to compute a minimal model can serve as a measure of semantic processing difficulty. Thus, especially comparing the first to the third example suggests that the latter should cause more difficulty.
APPENDIX B

Discourse representation theory (DRT)

In this appendix a number of key concepts of DRT will be introduced (see also Kamp & Reyle (1993) and the very nice introduction by Blackburn & Bos (1999) from whom much is borrowed here). Emphasis will be on the representation language DRT employs, a language based on box-like structures called DRSs. The construction of a DRS starts with the syntactic representation and transforms it piece by piece into a semantic representation corresponding to a mental model constructed during the process of discourse comprehension. In the system used in this dissertation the mental (minimal) model is computed using EC but DRT is nevertheless very important because it provides a disambiguated discourse representation that can be build in a word-by-word fashion. It takes care of peculiarities that have to do with discourse, the domain for which DRT was invented.

Most messages in natural language consist of sequences of several sentences, they are discourses. The basic question of DRT is how the meaning of a discourse can be represented semantically. Discourse meaning cannot be captured by simply conjoining sentences. Consider the following toy discourse in (1).

(1) A man sleeps. He snorts.

The meaning of this discourse isn't captured by formula (2).

(2) \( \exists x (\text{man}(x) \land \text{sleeps}(x)) \land \text{snorts}(y) \)

Even if some postprocessing guaranteed that the free variable \( y \) is replaced by \( x \) yielding (3), an important aspect of the discourse meaning is missing.

(3) \( \exists x (\text{man}(x) \land \text{sleeps}(x) \land \text{snorts}(x)) \)

That is, discourse meaning should also encode how the discourse actually works. Intuitively, \( \text{a man sleeps} \) sets the stage. A new discourse referent is introduced. He is a man and he sleeps. The anaphor in the second sentence refers back to this discourse referent and attributes a new property to him, he is snoring. What the first-order formula in (3) doesn't encode is the contextual relevance of the discourse referent. This missing piece of information is the so called Context Change Potential, the way an utterance changes the context in which subsequent utterances have to be interpreted. DRT models context change potential in a natural way by incrementally constructing the context as
the discourse moves along. I will show informally how this works by interpreting the previous discourse step by step using the (standard) DRS construction algorithm. But first we need a definition of what a DRS is.

Definition 17 DRSs and conditions:

1. If \( x_1, \ldots, x_n \) are discourse referents and \( \varphi_1, \ldots, \varphi_m \) are conditions then

\[
(x_1, \ldots, x_n; \varphi_1; \ldots; \varphi_m)
\]

is a DRS.

2. If \( P \) is a predicate of arity \( n \) and \( \tau_1, \ldots, \tau_n \) are terms (= constants or discourse referents), then \( P(\tau_1, \ldots, \tau_n) \) is a condition.

3. If \( \tau_1 \) and \( \tau_2 \) are terms, then \( \tau_1 = \tau_2 \) is a condition.

4. If \( B \) is a DRS, then \( \neg B \) is a condition.

5. If \( B_1 \) and \( B_2 \) are DRSs, then \( B_1 \lor B_2 \) is a condition.

6. If \( B_1 \) and \( B_2 \) are DRSs, then \( B_1 \Rightarrow B_2 \) is a condition.

7. Nothing else is a DRS or a condition

b.1 Processing a sample discourse

The standard DRS construction algorithm processes a discourse sentence by sentence with sentences being interpreted top-down, from left to right.¹ We start with the first sentence \( a \) man sleeps and with the empty DRS, that is, the DRS containing no prior information (A precise formulation of the DRS construction rules can be found in Kamp & Reyle (1993)).

We work our way down the tree top-down and left to right. First, we go from the S node to the NP node and further down to the Det node which tells us what kind of NP we are dealing with. It's an indefinite noun phrase and we need the according construction rule. Indefinites introduce a new discourse referent into the set of discourse referents, so we have to add a new referent \( x \).

Next, we delete the already processed Det subtree leaving the discourse referent \( x \) instead and move on to the sister node, the noun \( \text{man} \). It introduces the condition \( \text{man}(x) \) into the DRS.

Now, the whole NP has been processed and is replaced by the discourse referent \( x \) instead and move on to the sister node, the V phrase. It consists of the intransitive verb \( \text{sleeps} \). The algorithm instructs us to introduce a further constraint on the NP discourse referent, the one-place predicate \( \text{sleep}(x) \). Accordingly, we add the condition to the DRS.

The whole sentence has been processed leaving no unresolved syntactic structure to the construction algorithm. Now, the second sentence can be taken into account and will be interpreted within the context of the first. The new sentence S node accesses

1. This is just one out of many possible construction algorithms. It is, however, especially well suited for my purposes because processing from left to right is a necessary prerequisite for word-by-word processing.
both the set of discourse referents and the set of conditions. This makes it very easy to perform the required anaphora resolution.

First the NP has to be interpreted. The NP is a pronoun and the according construction rule states that a new discourse referent \( y \) and a condition \( y = ? \) have to be introduced. The question mark has to be identified with an accessible, earlier introduced discourse referent. DRT defines exactly how the accessibility relation works. It's a geometric concept which is defined in terms of how DRSs are nested one inside another. DRS \( B_1 \) accessible from DRS \( B_2 \) when \( B_1 \) equals \( B_2 \) or when \( B_1 \) subordinates \( B_2 \). In our example, there is only one discourse referent that has been introduced so far, namely discourse referent \( x \) and this discourse referent is accessible. Thus, we have to update the DRS as follows:

\[
\begin{align*}
\text{Step 5:} & \quad x \\
\text{man(x)} & \quad \text{sleep(x)} \\
\text{S} & \quad \text{IV} \\
\text{snorts} & \quad \text{he} \\
\end{align*}
\]

Step 6 finishes the NP subtree. Again we have to move up and to the left and interpret the VP subtree. We do this exactly the same way like in the preceding sentence and obtain the final DRS:

\[
\begin{align*}
\text{Step 6:} & \quad x,y \\
\text{man(x)} & \quad \text{sleep(x)} \\
\text{y = x} & \quad \text{Pro} \\
\text{S} & \quad \text{IV} \\
\text{snorts} & \quad \text{he} \\
\end{align*}
\]

This DRS captures both the truth conditions and the context change potential of the discourse appropriately. Discourse entities are entered into the discourse model as the discourse proceeds and anaphoric links have to be established between the sentences by identifying discourse referents of the anaphor with its antecedent.

When is a DRS satisfied? One standard way to answer this question is to use an embedding semantics. The basic intuition behind an embedding semantics is that a DRS is a picture of the part of the world which is relevant for interpreting the discourse. A DRS is satisfied with respect to a model if and only if it is possible to associate the discourse referents with entities in the model in a way that they fulfill the conditions of the DRS. However, the model theoretic semantics of DRT isn't needed in the combined DRT/EC framework because DRSs only provide the input to EC which actually computes discourse models. The way the framework works, a DRS gives rise to the computation of a unique, minimal model. Thus, DRSs the way they are used here serve as (unambiguous) construction plans for mental models in a very concrete sense.
### Target sentences in Experiment 1

**Table C.1.** Items used in Experiment 1

<table>
<thead>
<tr>
<th>Item</th>
<th>Condition</th>
<th>Sentence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Ach./for</td>
<td>Der Athlet gewann/den Marathon/zwei Stunden lang/bei/der letzten Olympiade.</td>
</tr>
<tr>
<td></td>
<td>Ach./in</td>
<td>Der Athlet gewann/den Marathon/in zwei Stunden/bei/der letzten Olympiade.</td>
</tr>
<tr>
<td></td>
<td>Acc./for</td>
<td>Der Athlet absolvierte/den Marathon/zwei Stunden lang/bei/der letzten Olympiade.</td>
</tr>
<tr>
<td></td>
<td>Acc./in</td>
<td>Der Athlet absolvierte/den Marathon/in zwei Stunden/bei/der letzten Olympiade.</td>
</tr>
<tr>
<td>2</td>
<td>Ach./for</td>
<td>Der Bergsteiger erreichte/den Gipfel/eine halbe Stunde lang/trotz/ eines schweren Unwetters.</td>
</tr>
<tr>
<td></td>
<td>Ach./in</td>
<td>Der Bergsteiger erreichte/den Gipfel/in einer halben Stunde/trotz/ eines schweren Unwetters.</td>
</tr>
<tr>
<td></td>
<td>Acc./for</td>
<td>Der Bergsteiger bestieg/den Gipfel/eine halbe Stunde lang/trotz/ eines schweren Unwetters.</td>
</tr>
<tr>
<td></td>
<td>Acc./in</td>
<td>Der Bergsteiger bestieg/den Gipfel/in einer halben Stunde/trotz/ eines schweren Unwetters.</td>
</tr>
<tr>
<td>3</td>
<td>Ach./for</td>
<td>Die Armee eroberte/die Stadt/sieben Tage lang/während/des dreißigjährigen Kriegs.</td>
</tr>
<tr>
<td></td>
<td>Ach./in</td>
<td>Die Armee eroberte/die Stadt/in sieben Tagen/während/des dreißigjährigen Kriegs.</td>
</tr>
<tr>
<td></td>
<td>Acc./for</td>
<td>Die Armee zerstörte/die Stadt/sieben Tage lang/während/des dreißigjährigen Kriegs.</td>
</tr>
<tr>
<td></td>
<td>Acc./in</td>
<td>Die Armee zerstörte/die Stadt/in sieben Tagen/während/des dreißigjährigen Kriegs.</td>
</tr>
<tr>
<td>4</td>
<td>Ach./for</td>
<td>Die Mannschaft verlor/den Wettkampf/eine halbe Stunde lang/beim/letztjährigen Turnfest.</td>
</tr>
<tr>
<td></td>
<td>Ach./in</td>
<td>Die Mannschaft verlor/den Wettkampf/in einer halben Stunde/beim/letztjährigen Turnfest.</td>
</tr>
<tr>
<td></td>
<td>Acc./for</td>
<td>Die Mannschaft bestritt/den Wettkampf/eine halbe Stunde lang/beim/letztjährigen Turnfest.</td>
</tr>
<tr>
<td></td>
<td>Acc./in</td>
<td>Die Mannschaft bestritt/den Wettkampf/in einer halben Stunde/beim/letztjährigen Turnfest.</td>
</tr>
<tr>
<td>item</td>
<td>condition</td>
<td>sentence</td>
</tr>
<tr>
<td>------</td>
<td>-----------</td>
<td>---------</td>
</tr>
<tr>
<td>5.5</td>
<td>Ach./for</td>
<td>Der Psychologe/verstand/das Problem/eine Sitzung lang/bei/einer Eheberatung.</td>
</tr>
<tr>
<td></td>
<td>Ach./in</td>
<td>Der Psychologe/verstand/das Problem/in einer Sitzung/bei/einer Eheberatung.</td>
</tr>
<tr>
<td></td>
<td>Acc./for</td>
<td>Der Psychologe/erfuhr/das Problem/eine Sitzung lang/bei/einer Eheberatung.</td>
</tr>
<tr>
<td></td>
<td>Acc./in</td>
<td>Der Psychologe/erfuhr/das Problem/in einer Sitzung/bei/einer Eheberatung.</td>
</tr>
<tr>
<td>6.5</td>
<td>Ach./for</td>
<td>Der Kommissar/ertappte/den Dieb/eine Stunde lang/beim/Juwelenraub im Juweliergeschäft.</td>
</tr>
<tr>
<td></td>
<td>Ach./in</td>
<td>Der Kommissar/ertappte/den Dieb/in einer Stunde/beim/Juwelenraub im Juweliergeschäft.</td>
</tr>
<tr>
<td></td>
<td>Acc./for</td>
<td>Der Kommissar/überführte/den Dieb/eine Stunde lang/beim/Juwelenraub im Juweliergeschäft.</td>
</tr>
<tr>
<td></td>
<td>Acc./in</td>
<td>Der Kommissar/überführte/den Dieb/in einer Stunde/beim/Juwelenraub im Juweliergeschäft.</td>
</tr>
<tr>
<td>8.5</td>
<td>Ach./for</td>
<td>Der Jäger/entdeckte/die Falle/eine halbe Stunde lang/im/Unterholz.</td>
</tr>
<tr>
<td></td>
<td>Ach./in</td>
<td>Der Jäger/entdeckte/die Falle/in einer halben Stunde/im/Unterholz.</td>
</tr>
<tr>
<td></td>
<td>Acc./for</td>
<td>Der Jäger/errichtete/die Falle/eine halbe Stunde lang/im/Unterholz.</td>
</tr>
<tr>
<td></td>
<td>Acc./in</td>
<td>Der Jäger/errichtete/die Falle/in einer halben Stunde/im/Unterholz.</td>
</tr>
<tr>
<td>9.5</td>
<td>Ach./for</td>
<td>Der Schneider/erspähte/den Rock/zwei Minuten lang/auf/seinem Arbeitstisch.</td>
</tr>
<tr>
<td></td>
<td>Ach./in</td>
<td>Der Schneider/erspähte/den Rock/in zwei Minuten/auf/seinem Arbeitstisch.</td>
</tr>
<tr>
<td></td>
<td>Acc./for</td>
<td>Der Schneider/weitete/den Rock/zwei Minuten lang/auf/seinem Arbeitstisch.</td>
</tr>
<tr>
<td></td>
<td>Acc./in</td>
<td>Der Schneider/weitete/den Rock/in zwei Minuten/auf/seinem Arbeitstisch.</td>
</tr>
<tr>
<td>10.5</td>
<td>Ach./for</td>
<td>Der Metzger/erlegte/das Schwein/eine halbe Stunde lang/im/Tübinger Schlachthof.</td>
</tr>
<tr>
<td></td>
<td>Ach./in</td>
<td>Der Metzger/erlegte/das Schwein/in einer halben Stunde/im/Tübinger Schlachthof.</td>
</tr>
<tr>
<td></td>
<td>Acc./for</td>
<td>Der Metzger/zerlegte/das Schwein/eine halbe Stunde lang/im/Tübinger Schlachthof.</td>
</tr>
<tr>
<td></td>
<td>Acc./in</td>
<td>Der Metzger/zerlegte/das Schwein/in einer halben Stunde/im/Tübinger Schlachthof.</td>
</tr>
</tbody>
</table>

Appendix C. Target sentences in Experiment 1

<table>
<thead>
<tr>
<th>item</th>
<th>condition</th>
<th>sentence</th>
</tr>
</thead>
<tbody>
<tr>
<td>11.5</td>
<td>Ach./for</td>
<td>Der Diplomand/erhielt/das Zeugnis/dreißig Minuten lang/trotz/großen Andrangs im Dekanat.</td>
</tr>
<tr>
<td></td>
<td>Ach./in</td>
<td>Der Diplomand/erhielt/das Zeugnis/in dreißig Minuten/trotz/großen Andrangs im Dekanat.</td>
</tr>
<tr>
<td></td>
<td>Acc./for</td>
<td>Der Diplomand/beantragte/das Zeugnis/dreißig Minuten lang/trotz/großen Andrangs im Dekanat.</td>
</tr>
<tr>
<td></td>
<td>Acc./in</td>
<td>Der Diplomand/beantragte/das Zeugnis/in dreißig Minuten/trotz/großen Andrangs im Dekanat.</td>
</tr>
<tr>
<td>12.5</td>
<td>Ach./for</td>
<td>Der Bauarbeiter/fand/den Schubkarren/fünf Minuten lang/vor/dem Lagerraum.</td>
</tr>
<tr>
<td></td>
<td>Ach./in</td>
<td>Der Bauarbeiter/fand/den Schubkarren/in fünf Minuten/vor/dem Lagerraum.</td>
</tr>
<tr>
<td></td>
<td>Acc./for</td>
<td>Der Bauarbeiter/entlud/den Schubkarren/fünf Minuten lang/vor/dem Lagerraum.</td>
</tr>
<tr>
<td></td>
<td>Acc./in</td>
<td>Der Bauarbeiter/entlud/den Schubkarren/in fünf Minuten/vor/dem Lagerraum.</td>
</tr>
<tr>
<td>13.5</td>
<td>Ach./for</td>
<td>Der Student/verstand/den Aufsatz/zwei Tage lang/bei/der Vorbereitung des Seminars.</td>
</tr>
<tr>
<td></td>
<td>Ach./in</td>
<td>Der Student/verstand/den Aufsatz/in zwei Tagen/bei/der Vorbereitung des Seminars.</td>
</tr>
<tr>
<td></td>
<td>Acc./for</td>
<td>Der Student/verfasste/den Aufsatz/zwei Tage lang/bei/der Vorbereitung des Seminars.</td>
</tr>
<tr>
<td></td>
<td>Acc./in</td>
<td>Der Student/verfasste/den Aufsatz/in zwei Tagen/bei/der Vorbereitung des Seminars.</td>
</tr>
<tr>
<td>14.5</td>
<td>Ach./for</td>
<td>Der Ingenieur/erkannte/den Fehler/eine Woche lang/bei/der Erprobung des Prototyps.</td>
</tr>
<tr>
<td></td>
<td>Ach./in</td>
<td>Der Ingenieur/erkannte/den Fehler/in einer Woche/bei/der Erprobung des Prototyps.</td>
</tr>
<tr>
<td></td>
<td>Acc./for</td>
<td>Der Ingenieur/behob/den Fehler/eine Woche lang/bei/der Erprobung des Prototyps.</td>
</tr>
<tr>
<td></td>
<td>Acc./in</td>
<td>Der Ingenieur/behob/den Fehler/in einer Woche/bei/der Erprobung des Prototyps.</td>
</tr>
<tr>
<td>15.5</td>
<td>Ach./for</td>
<td>Der Tüftler/erfand/die Maschine/fünf Tage lang/in/seiner unaufgeräumten Werkstatt.</td>
</tr>
<tr>
<td></td>
<td>Ach./in</td>
<td>Der Tüftler/erfand/die Maschine/in fünf Tagen/in/seiner unaufgeräumten Werkstatt.</td>
</tr>
<tr>
<td></td>
<td>Acc./for</td>
<td>Der Tüftler/erfand/die Maschine/fünf Tage lang/in/seiner unaufgeräumten Werkstatt.</td>
</tr>
<tr>
<td></td>
<td>Acc./in</td>
<td>Der Tüftler/erfand/die Maschine/in fünf Tagen/in/seiner unaufgeräumten Werkstatt.</td>
</tr>
<tr>
<td>16.5</td>
<td>Ach./for</td>
<td>Der Manager/empfing/die Nachricht/fünf Minuten lang/in/Form einer E-Mail.</td>
</tr>
<tr>
<td></td>
<td>Ach./in</td>
<td>Der Manager/empfing/die Nachricht/in fünf Minuten/in/Form einer E-Mail.</td>
</tr>
</tbody>
</table>
### The Processing of Events

<table>
<thead>
<tr>
<th>item</th>
<th>condition</th>
<th>sentence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acc./for</td>
<td>Der Manager/verbreitete/die Nachricht/fünf Minuten lang/in/Form einer E-Mail.</td>
<td>Der Manager/verbreitete/die Nachricht/fünf Minuten lang/in/Form einer E-Mail.</td>
</tr>
<tr>
<td>Acc./in</td>
<td>Der Manager/verbreitete/die Nachricht/in fünf Minuten/in/Form einer E-Mail.</td>
<td>Der Manager/verbreitete/die Nachricht/in fünf Minuten/in/Form einer E-Mail.</td>
</tr>
<tr>
<td>17</td>
<td>Ach./for</td>
<td>Der Klempner/bemerkte/den Wasserrohrbruch/dreißig Minuten lang/in/der alten Mietskaserne.</td>
</tr>
<tr>
<td>Ach./in</td>
<td>Der Klempner/bemerkte/den Wasserrohrbruch/in dreißig Minuten/in/der alten Mietskaserne.</td>
<td></td>
</tr>
<tr>
<td>Acc./for</td>
<td>Der Klempner/reparierte/den Wasserrohrbruch/dreißig Minuten lang/in/der alten Mietskaserne.</td>
<td></td>
</tr>
<tr>
<td>Acc./in</td>
<td>Der Klempner/reparierte/den Wasserrohrbruch/in dreißig Minuten/in/der alten Mietskaserne.</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>Ach./for</td>
<td>Der Höhlenforscher/verließ/die Höhle/fünfzehn Minuten lang/nach/deren Entdeckung.</td>
</tr>
<tr>
<td>Ach./in</td>
<td>Der Höhlenforscher/verließ/die Höhle/in fünfzehn Minuten/nach/deren Entdeckung.</td>
<td></td>
</tr>
<tr>
<td>Acc./for</td>
<td>Der Höhlenforscher/durchquerte/die Höhle/fünfzehn Minuten lang/nach/deren Entdeckung.</td>
<td></td>
</tr>
<tr>
<td>Acc./in</td>
<td>Der Höhlenforscher/durchquerte/die Höhle/in fünfzehn Minuten/nach/deren Entdeckung.</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>Ach./for</td>
<td>Der Indianer/bewältigte/den Fluss/drei Minuten lang/trotz/der reißenden Strömung.</td>
</tr>
<tr>
<td>Ach./in</td>
<td>Der Indianer/bewältigte/den Fluss/in drei Minuten/trotz/der reißenden Strömung.</td>
<td></td>
</tr>
<tr>
<td>Acc./for</td>
<td>Der Indianer/durchschwamm/den Fluss/drei Minuten lang/trotz/der reißenden Strömung.</td>
<td></td>
</tr>
<tr>
<td>Acc./in</td>
<td>Der Indianer/durchschwamm/den Fluss/in drei Minuten/trotz/der reißenden Strömung.</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>Ach./for</td>
<td>Der Mathematiker/durschaute/das Theorem/zwei Tage lang/trotz/einer schweren Grippe.</td>
</tr>
<tr>
<td>Ach./in</td>
<td>Der Mathematiker/durschaute/das Theorem/in zwei Tagen/trotz/einer schweren Grippe.</td>
<td></td>
</tr>
<tr>
<td>Acc./for</td>
<td>Der Mathematiker/bewies/das Theorem/zwei Tage lang/trotz/einer schweren Grippe.</td>
<td></td>
</tr>
<tr>
<td>Acc./in</td>
<td>Der Mathematiker/bewies/das Theorem/in zwei Tagen/trotz/einer schweren Grippe.</td>
<td></td>
</tr>
</tbody>
</table>

### APPENDIX D

**Lexical frequencies of verbs in Experiment 2**

<table>
<thead>
<tr>
<th>verb</th>
<th>activity</th>
<th>accomplishment</th>
<th>achievement</th>
<th>state</th>
</tr>
</thead>
<tbody>
<tr>
<td>backen</td>
<td>61</td>
<td>20</td>
<td>0</td>
<td>19</td>
</tr>
<tr>
<td>baden</td>
<td>79</td>
<td>3</td>
<td>0</td>
<td>18</td>
</tr>
<tr>
<td>faseln</td>
<td>92</td>
<td>8</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>flanieren</td>
<td>95</td>
<td>3</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>geistern</td>
<td>98</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>joggen</td>
<td>84</td>
<td>13</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>knutschen</td>
<td>80</td>
<td>7</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td>marschieren</td>
<td>58</td>
<td>40</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>musizieren</td>
<td>86</td>
<td>11</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>paddeln</td>
<td>52</td>
<td>24</td>
<td>24</td>
<td>0</td>
</tr>
<tr>
<td>putzen</td>
<td>43</td>
<td>42</td>
<td>14</td>
<td>1</td>
</tr>
<tr>
<td>radeln</td>
<td>54</td>
<td>37</td>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td>rennen</td>
<td>53</td>
<td>29</td>
<td>18</td>
<td>0</td>
</tr>
<tr>
<td>rudern</td>
<td>68</td>
<td>15</td>
<td>16</td>
<td>1</td>
</tr>
<tr>
<td>schimpfen</td>
<td>86</td>
<td>3</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td>schwimmen</td>
<td>56</td>
<td>21</td>
<td>22</td>
<td>1</td>
</tr>
<tr>
<td>streiten</td>
<td>96</td>
<td>1</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>surfen</td>
<td>80</td>
<td>7</td>
<td>2</td>
<td>11</td>
</tr>
<tr>
<td>tanzen</td>
<td>67</td>
<td>5</td>
<td>4</td>
<td>24</td>
</tr>
<tr>
<td>tratschen</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>bauen</td>
<td>32</td>
<td>57</td>
<td>8</td>
<td>3</td>
</tr>
<tr>
<td>designen</td>
<td>28</td>
<td>70</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>entladen</td>
<td>8</td>
<td>91</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>entschlüsseln</td>
<td>5</td>
<td>77</td>
<td>11</td>
<td>7</td>
</tr>
</tbody>
</table>
APPENDIX E

Target sentences in Experiment 2

Table E.1. Items used in Experiment 2

<table>
<thead>
<tr>
<th>item</th>
<th>condition</th>
<th>sentence</th>
</tr>
</thead>
</table>
Hannes entwarf eine halbe Stunde lang eine neue Skizze.

Helga entschlüsselte in mehreren Stunden die geheime Botschaft.

Markus entlud vierzig Minuten lang im Supermarkt Paletten.

Anna entwickelte in drei Stunden großformatige Photos.

Anna entwickelte drei Stunden lang ein großformatiges Photo.

Bernd musizierte eine halbe Stunde lang auf der Gitarre.

Katrin paddelte drei Tage lang auf dem offenen Meer.

Anna entwickelte in drei Stunden ein großformatiges Foto.

Katrin konstruierte drei Tage lang für die neue Ausstellung Skulpturen.

Stefanie putzte den ganzen Nachmittag total verdrecktes Geschirr.
<table>
<thead>
<tr>
<th>item</th>
<th>condition</th>
<th>sentence</th>
</tr>
</thead>
<tbody>
<tr>
<td>28</td>
<td>act./for/act.-cont.</td>
<td>Holger/strich/eine Viertelstunde lang/nahe am Regal.</td>
</tr>
<tr>
<td></td>
<td>act./for/act.-cont.</td>
<td>Holger/strich/eine Viertelstunde lang/nahe am Regal.</td>
</tr>
<tr>
<td></td>
<td>act./for/act.-cont.</td>
<td>Holger/strich/eine Viertelstunde lang/nahe am Regal.</td>
</tr>
<tr>
<td></td>
<td>act./for/act.-cont.</td>
<td>Holger/strich/eine Viertelstunde lang/nahe am Regal.</td>
</tr>
</tbody>
</table>
### The Processing of Events

#### APPENDIX F

**Discourses in Experiment 3**

<table>
<thead>
<tr>
<th>Table F.1. Items used in Experiment 3</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>item</th>
<th>condition</th>
<th>sentence</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>acc./in/act.-cont.</td>
<td>Franz/skizzierte/in einem Vormittag/äußerst innovative/ Entwürfe.</td>
</tr>
<tr>
<td></td>
<td>acc./in/acc.-cont.</td>
<td>Franz/skizzierte/in einem Vormittag/einen äußerst/ innovativen Entwurf.</td>
</tr>
<tr>
<td>40</td>
<td>acc./for/act.-cont.</td>
<td>Franziska/zerlegte/zwei Stunden lang/fangfrische/Forellen.</td>
</tr>
<tr>
<td></td>
<td>acc./in/act.-cont.</td>
<td>Franziska/zerlegte/in zwei Stunden/fangfrische/Forellen.</td>
</tr>
<tr>
<td></td>
<td>acc./for/acc.-cont.</td>
<td>Franziska/zerlegte/zwei Stunden lang/drei fangfrische/ Forellen.</td>
</tr>
<tr>
<td></td>
<td>acc./in/acc.-cont.</td>
<td>Franziska/zerlegte/in zwei Stunden/drei fangfrische/Forellen.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>item</th>
<th>condition</th>
<th>discourse</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>target-in</td>
<td>Heute morgen/schwamm/er/in nur dreiölf/Minenuten.</td>
</tr>
<tr>
<td></td>
<td>target-for</td>
<td>Heute morgen/schwamm/er/dreiölf/Minenuten/lang.</td>
</tr>
<tr>
<td>2</td>
<td>telic context</td>
<td>Seit dem letzten Sommer/joggt/Maria/jeden Tag/einmal/durch den ganzen Park. Als sie/damit/begann,/brauchte sie mehr als eineinhalbe Stunden dafür,/aber/jetzt/läuft sie/viel schneller.</td>
</tr>
<tr>
<td></td>
<td>atelic context</td>
<td>Seit dem letzten Sommer/joggt/Maria/jeden Tag/einmal/im nahe gelegenen Park. Als sie/damit/begann,/tat ihr noch alles weh,/aber/mittlerweile/macht es ihr/richtig Spaß.</td>
</tr>
<tr>
<td></td>
<td>target-in</td>
<td>Heute Mittag/joggte/sie/in nur einer/Stunde.</td>
</tr>
<tr>
<td></td>
<td>target-for</td>
<td>Heute Mittag/joggte/sie/eine Stunde/lang.</td>
</tr>
<tr>
<td>3</td>
<td>telic context</td>
<td>Markus/trainiert/auf ein Radrennen/und radelt deshalb/jeden Tag/sechzig Kilometer. Vor sechs Wochen/brauchte er dafür/noch gute zwei Stunden,/aber/er hat sich/deutlich/gesteigert.</td>
</tr>
<tr>
<td></td>
<td>atelic context</td>
<td>Markus/trainiert/auf dem Rennrad/und radelt/öfters/auf der Alb herum. Vor sechs Wochen/kam er/kaum die Steige hinauf,/aber/mittlerweile/strengt es ihn/kaum noch an.</td>
</tr>
<tr>
<td></td>
<td>target-in</td>
<td>Heute Nachmittag/radelte/er/in nur eineinhalb/Stunden.</td>
</tr>
<tr>
<td></td>
<td>target-for</td>
<td>Heute Nachmittag/radelte/er/eineinhalb Stunden/lang.</td>
</tr>
<tr>
<td>4</td>
<td>telic context</td>
<td>Dieter/ist ein begeisterter Läufer/und läuft/jeden Morgen/eine Strecke/von zehn Kilometern. Als er/damit/angefangen hat,/brauchte er noch eine Stunde/ dafür/jetzt ist er/sehr viel schneller.</td>
</tr>
</tbody>
</table>
### Appendix F. Discourses in Experiment 3

<table>
<thead>
<tr>
<th>item</th>
<th>condition</th>
<th>discourse</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>telic</td>
<td>Seit mehreren Jahren/wandert/Martin/jedes Jahr/einmal/quer über die Alpen. Im ersten Jahr/benötigte er/über zwei Wochen für die Distanz,/aber/mittlerweile/geht es/deutlich besser.</td>
</tr>
<tr>
<td></td>
<td>target-in</td>
<td>Letzte Woche/marschierte/er/in nur vier/Stunden.</td>
</tr>
<tr>
<td></td>
<td>target-for</td>
<td>Gestern früh/marschierte/er/in nur vier/Stunden/lang.</td>
</tr>
<tr>
<td>6</td>
<td>telic</td>
<td>Kerstin/rudert/seit/diesem Frühjahr/jeden Morgen/zehn Kilometer. Anfangs/benötigte sie/noch/über eine Stunde für die Distanz,/aber/sie wird von Tag zu Tag/schneller.</td>
</tr>
<tr>
<td></td>
<td>target-in</td>
<td>Vorgestern Abend/putzte/er/in nur zwei/Stunden.</td>
</tr>
<tr>
<td></td>
<td>target-for</td>
<td>Vorgestern Abend/putzte/er/zwei Stunden/lang.</td>
</tr>
<tr>
<td>7</td>
<td>telic</td>
<td>Letzten Samstag/paddelte/sie/in nur einer halben/Stunde.</td>
</tr>
<tr>
<td></td>
<td>atelic</td>
<td>Letzten Samstag/paddelte/sie/in eine Stunde/lang.</td>
</tr>
<tr>
<td></td>
<td>target-in</td>
<td>Gestern früh/Ruderte/sie/in nur einer halben/Stunde/lang.</td>
</tr>
<tr>
<td></td>
<td>target-for</td>
<td>Gestern früh/Ruderte/sie/eine halbe Stunde/lang.</td>
</tr>
<tr>
<td>8</td>
<td>telic</td>
<td>Diese Sommer/wanderte/er/in nur einer halben/Stunde.</td>
</tr>
<tr>
<td></td>
<td>atelic</td>
<td>Diese Sommer/wanderte/er/eine Woche/lang.</td>
</tr>
<tr>
<td></td>
<td>target-in</td>
<td>Gestern früh/DDerte/sie/in nur einer halben/Stunde/lang.</td>
</tr>
<tr>
<td></td>
<td>target-for</td>
<td>Gestern früh/DDerte/sie/eine halbe Stunde/lang.</td>
</tr>
<tr>
<td>9</td>
<td>telic</td>
<td>Gerd/marschiert/bei der Bundeswehr/jede Woche/fünfundzwanzig Kilometer. Vor fünf Wochen/brauchte er/dafür/noch fast einen Tag,/mittlerweile/ist er/viel schneller.</td>
</tr>
<tr>
<td></td>
<td>atelic</td>
<td>Gerd/marschiert/bei der Bundeswehr/jede Woche/mit schwerem Gepäck. Vor fünf Wochen/brach er/noch/fast unter der Last zusammen,/mittlerweile/geht es/jedoch/immer besser.</td>
</tr>
<tr>
<td></td>
<td>target-in</td>
<td>Letzte Woche/marschierte/er/in nur vier/Stunden.</td>
</tr>
<tr>
<td></td>
<td>target-for</td>
<td>Letzte Woche/marschierte/er/zwei Stunden/lang.</td>
</tr>
<tr>
<td></td>
<td>target-in</td>
<td>Heute Morgen/kraulzte/er/in nur fünfundvierzig Minuten.</td>
</tr>
<tr>
<td></td>
<td>target-for</td>
<td>Heute Morgen/kraulzte/er/fünfundvierzig Minuten/lang.</td>
</tr>
<tr>
<td>11</td>
<td>telic</td>
<td>Um sich das Rauchen/abzuwöhnen/rennt/Maria/jeden Morgen/fünf Kilometer weit. Zu Beginn/brauchte sie/noch sehr lange dafür,/aber sie wird/von Tag zu Tag/schneller.</td>
</tr>
<tr>
<td></td>
<td>atelic</td>
<td>Um sich das Rauchen/abzuwöhnen/rennt/Maria/jeden Morgen/im Wald. Zu Beginn/brachte sie/nach fünf Minuten/hustend zusammen,/aber es geht/von Tag zu Tag/besser.</td>
</tr>
<tr>
<td></td>
<td>target-in</td>
<td>Gestern früh/rannste/sie/in nur einer halben/Stunde/lang.</td>
</tr>
<tr>
<td></td>
<td>target-for</td>
<td>Gestern früh/rannste/sie/eine halbe Stunde/lang.</td>
</tr>
<tr>
<td>item</td>
<td>condition</td>
<td>discourse</td>
</tr>
<tr>
<td>------</td>
<td>-----------</td>
<td>-----------</td>
</tr>
<tr>
<td>13</td>
<td>teile context</td>
<td>Für die Ausbildung zur Sekretärin tippt Julia jeden Morgen. Anfangs benötigte sie noch mehr als drei Stunden dafür, aber mittlerweile ist sie schon viel schneller.</td>
</tr>
<tr>
<td></td>
<td>ateleic context</td>
<td>Für die Ausbildung zur Sekretärin tippt Julia jeden Morgen auf dem Computer. Anfangs verlor sie bereits nach kurzer Zeit die Konzentration, aber mittlerweile geht es viel besser.</td>
</tr>
<tr>
<td></td>
<td>target-in</td>
<td>Heute morgen tippte sie in einen halben Stunden.</td>
</tr>
<tr>
<td></td>
<td>target-for</td>
<td>Heute morgen tippte sie in einen halben Stunden lang.</td>
</tr>
<tr>
<td>14</td>
<td>teile context</td>
<td>Herr Meier jätet seit zwei Jahren jeden Samstag das Unkraut im ganzen Vorgarten. Anfangs brauchte er noch recht lang dafür, aber mittlerweile hat er ein ausgeklügeltes System.</td>
</tr>
<tr>
<td></td>
<td>ateleic context</td>
<td>Herr Meier jätet seit zwei Jahren jeden Samstag Unkraut im Vorgarten. Anfangs verbrachte er noch ganze Tage im Garten, aber mittlerweile hat er keine rechte Lust mehr.</td>
</tr>
<tr>
<td></td>
<td>target-in</td>
<td>Diesen Samstag jätete er in nur einer Stunde.</td>
</tr>
<tr>
<td></td>
<td>target-for</td>
<td>Diesen Samstag jätete er in einer halben Stunden lang.</td>
</tr>
<tr>
<td>15</td>
<td>teile context</td>
<td>Seit Fritz in die Schule gekommen ist, liest er seiner Mutter jeden Tag ein Märchen vor. Zu Beginn musste er sich noch sehr anstrengen und brauchte länger als eine Stunde dafür.</td>
</tr>
<tr>
<td></td>
<td>ateleic context</td>
<td>Seit Fritz in die Schule gekommen ist, liest er seiner Mutter jeden Tag in seinem Märchenbuch. Zu Beginn wurde er bereits nach wenigen Minuten müde, aber heutzutage liest er viel besser.</td>
</tr>
<tr>
<td></td>
<td>target-in</td>
<td>Gestern las er in nur zwanzig Minuten.</td>
</tr>
<tr>
<td></td>
<td>target-for</td>
<td>Gestern las er in zwanzig Minuten lang.</td>
</tr>
<tr>
<td>16</td>
<td>teile context</td>
<td>Seit Klaus eine Freundin hat, kocht er jede Woche ein aufwändiges Abendessen für sie. Anfangs brauchte er für ein Abendessen noch zwei Stunden, aber mittlerweile ist er routiniert.</td>
</tr>
<tr>
<td></td>
<td>ateleic context</td>
<td>Seit Klaus eine neue Freundin hat, kocht er sehr häufig für sie. Anfangs kocht er nur Kleinigkeiten, aber neuerdings wagt er sich auch an schwierige Rezepte.</td>
</tr>
<tr>
<td></td>
<td>target-in</td>
<td>Gestern Abend kocht er in nur einer Stunde.</td>
</tr>
<tr>
<td></td>
<td>target-for</td>
<td>Gestern Abend kocht er in einer halben Stunde lang.</td>
</tr>
<tr>
<td></td>
<td>ateleic context</td>
<td>Herr Bauer fegt seit Jahren jeden Samstag pünktlich um neun vor dem Haus. Zu Beginn verlor er sehr schnell die Lust, aber mittlerweile findet er es richtig entspannend.</td>
</tr>
<tr>
<td></td>
<td>target-in</td>
<td>Diesen Samstag legte er in nur zwei Stunden.</td>
</tr>
<tr>
<td></td>
<td>target-for</td>
<td>Diesen Samstag legte er in zwei Stunden lang.</td>
</tr>
<tr>
<td>18</td>
<td>teile context</td>
<td>Martina sprintet seit Jahren im Leichtathletiktraining jeden Montag zwei hundert Meter. Als sie damit begann brauchte sie über dreißig Sekunden dafür, aber sie ist nun viel schneller.</td>
</tr>
<tr>
<td></td>
<td>ateleic context</td>
<td>Martina sprintet seit Jahren im Leichtathletiktraining jeden Montag auf der Tartanbahn. Als sie damit begann konnte sie nur sehr kurz die nötige Geschwindigkeit halten.</td>
</tr>
<tr>
<td></td>
<td>target-in</td>
<td>Letzten Montag sprintete sie in nur zwanzig Sekunden.</td>
</tr>
<tr>
<td></td>
<td>target-for</td>
<td>Letzten Montag sprintete sie in zwanzig Sekunden lang.</td>
</tr>
<tr>
<td>19</td>
<td>teile context</td>
<td>Christine arbeitet seit diesem Jahr als Übersetzerin und übersetzt Gesetzes texte. Zu Beginn benötigte sie noch zwei Tage pro Text, aber sie wird immer schneller.</td>
</tr>
<tr>
<td></td>
<td>ateleic context</td>
<td>Christine arbeitet seit diesem Jahr als Übersetzerin vom Deutschen ins Japanische. Anfangs brauchte sie noch rasch eine Pause, aber mittlerweile hat sie viel Übung.</td>
</tr>
<tr>
<td></td>
<td>target-in</td>
<td>Heute übersetzte sie in nur fünf Stunden.</td>
</tr>
<tr>
<td></td>
<td>target-for</td>
<td>Heute übersetzte sie in fünf Stunden lang.</td>
</tr>
<tr>
<td>20</td>
<td>teile context</td>
<td>Seit Peter bei seinen Eltern ausgezogen ist, bügelt er jede Woche sein Lieblings hemd selbst. Zu Beginn brauchte er noch über eine Viertelstunde dafür, aber mittlerweile hat er Routine.</td>
</tr>
<tr>
<td></td>
<td>ateleic context</td>
<td>Seit Peter bei seinen Eltern ausgezogen ist, bügelt er selbst. Zu Beginn tat er es noch sehr widerwillig, aber mittlerweile macht es ihm nichts mehr aus.</td>
</tr>
<tr>
<td></td>
<td>target-in</td>
<td>Gestern bügelte er in nur fünf Minuten.</td>
</tr>
<tr>
<td></td>
<td>target-for</td>
<td>Gestern bügelte er in fünf Minuten lang.</td>
</tr>
<tr>
<td>21</td>
<td>teile context</td>
<td>Karsten trainiert seit letzten Jahr die zehntausend Meter. Einzelstrecke im Einschnellauf. Anfangs brauchte er noch über eine Stunde dafür, aber mittlerweile läuft er schon viel schneller.</td>
</tr>
<tr>
<td></td>
<td>ateleic context</td>
<td>Karsten läuft seit einigen Jahren jeden Winter regelmäßig Schlittschuh auf dem See. Anfangs holte er sich noch viele blaue Flecke, aber mittlerweile läuft er schon viel sicherer.</td>
</tr>
<tr>
<td></td>
<td>target-in</td>
<td>Heute morgen lief er in nur vierzig Minuten.</td>
</tr>
<tr>
<td></td>
<td>target-for</td>
<td>Heute morgen lief er in vierzig Minuten lang.</td>
</tr>
</tbody>
</table>
The Processing of Events

APPENDIX G

Target sentences in Experiment 4 (4a and 4b)

Table G.1. Items used in Experiment 4

<table>
<thead>
<tr>
<th>item (Exp.)</th>
<th>condition</th>
<th>sentence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (4a)</td>
<td>contr./SVOA</td>
<td>Die Athletin/gewann/die Goldmedaille/vor knapp drei Stunden/bei/den Para-Olympics.</td>
</tr>
<tr>
<td></td>
<td>coerc./SVOA</td>
<td>Die Athletin/gewann/die Goldmedaille/in knapp drei Stunden/bei/den Para-Olympics.</td>
</tr>
<tr>
<td></td>
<td>mism./SVOA</td>
<td>Die Athletin/gewann/die Goldmedaille/knapp drei Stunden lang/bei/den Para-Olympics.</td>
</tr>
<tr>
<td>1 (4b)</td>
<td>contr./SVOAO</td>
<td>Die Athletin/gewann/vor knapp drei Stunden/bei/den Para-Olympics/die Goldmedaille…</td>
</tr>
<tr>
<td></td>
<td>coerc./SVOAO</td>
<td>Die Athletin/gewann/in knapp drei Stunden/bei/den Para-Olympics/die Goldmedaille…</td>
</tr>
<tr>
<td></td>
<td>mism./SVOAO</td>
<td>Die Athletin/gewann/knapp drei Stunden lang/bei/den Para-Olympics/die Goldmedaille…</td>
</tr>
<tr>
<td>2 (4a)</td>
<td>contr./SVOA</td>
<td>Das Segelschiff/erreichte/den Hafen/vor einer halben Stunde/bei/stürmischer See.</td>
</tr>
<tr>
<td></td>
<td>coerc./SVOA</td>
<td>Das Segelschiff/erreichte/den Hafen/in einer halben Stunde/bei/stürmischer See.</td>
</tr>
<tr>
<td></td>
<td>mism./SVOA</td>
<td>Das Segelschiff/erreichte/den Hafen/eine halbe Stunde lang/bei/stürmischer See.</td>
</tr>
<tr>
<td>2 (4b)</td>
<td>contr./SVOAO</td>
<td>Das Segelschiff/erreichte/vor einer halben Stunde/bei/stürmischer See/den Hafen…</td>
</tr>
<tr>
<td></td>
<td>coerc./SVOAO</td>
<td>Das Segelschiff/erreichte/in einer halben Stunde/bei/stürmischer See/den Hafen…</td>
</tr>
<tr>
<td></td>
<td>mism./SVOAO</td>
<td>Das Segelschiff/erreichte/eine halbe Stunde lang/bei/stürmischer See/den Hafen…</td>
</tr>
<tr>
<td>3 (4a)</td>
<td>contr./SVOA</td>
<td>Die Armee/eroberte/die Stadt/vor sieben Tagen/trotz/heftigen Widerstands.</td>
</tr>
<tr>
<td></td>
<td>coerc./SVOA</td>
<td>Die Armee/eroberte/die Stadt/in sieben Tagen/trotz/heftigen Widerstands.</td>
</tr>
<tr>
<td></td>
<td>mism./SVOA</td>
<td>Die Armee/eroberte/die Stadt/sieben Tage lang/trotz/heftigen Widerstands.</td>
</tr>
<tr>
<td>Item (Exp.)</td>
<td>Condition</td>
<td>Sentence</td>
</tr>
<tr>
<td>------------</td>
<td>-----------</td>
<td>----------</td>
</tr>
<tr>
<td>3 (4b) contr./SAVO</td>
<td>Die Armee/eroberte/vor sieben Tagen/trotz/fehtigen Widerstands/die Stadt…</td>
<td></td>
</tr>
<tr>
<td></td>
<td>coerc./SAVO</td>
<td>Die Armee/eroberte/in sieben Tagen/trotz/fehtigen Widerstands/die Stadt…</td>
</tr>
<tr>
<td></td>
<td>mism./SAVO</td>
<td>Die Armee/eroberte/sieben Tage lang/trotz/fehtigen Widerstands/die Stadt…</td>
</tr>
<tr>
<td>4 (4a) contr./SAVO</td>
<td>Das Militär/exekutierte/den Regimegegner/vor einer Stunde/trotz/weltweiter Proteste/den Regimegegner…</td>
<td></td>
</tr>
<tr>
<td></td>
<td>coerc./SAVO</td>
<td>Das Militär/exekutierte/den Regimegegner/in einer Stunde/trotz/weltweiter Proteste/den Regimegegner…</td>
</tr>
<tr>
<td></td>
<td>mism./SAVO</td>
<td>Das Militär/exekutierte/den Regimegegner/eine Stunde lang/trotz/weltweiter Proteste/den Regimegegner…</td>
</tr>
<tr>
<td>4 (4b) contr./SAVO</td>
<td>Das Militär/exekutierte/vor einer Stunde/trotz/weltweiter Proteste/den Regimegegner…</td>
<td></td>
</tr>
<tr>
<td></td>
<td>coerc./SAVO</td>
<td>Das Militär/exekutierte/in einer Stunde/trotz/weltweiter Proteste/den Regimegegner…</td>
</tr>
<tr>
<td></td>
<td>mism./SAVO</td>
<td>Das Militär/exekutierte/eine Stunde lang/trotz/weltweiter Proteste/den Regimegegner…</td>
</tr>
<tr>
<td>5 (4a) contr./SAVO</td>
<td>Der Psychologe/erkannte/das Problem/vor einer halben Stunde/bei/einer Erheberung.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>coerc./SAVO</td>
<td>Der Psychologe/erkannte/das Problem/in einer halben Stunde/bei/einer Erheberung.</td>
</tr>
<tr>
<td></td>
<td>mism./SAVO</td>
<td>Der Psychologe/erkannte/das Problem/eine halbe Stunde lang/bei/einer Erheberung.</td>
</tr>
<tr>
<td>5 (4b) contr./SAVO</td>
<td>Der Psychologe/erkannte/vor einer halben Stunde/bei/einer Erheberung/das Problem…</td>
<td></td>
</tr>
<tr>
<td></td>
<td>coerc./SAVO</td>
<td>Der Psychologe/erkannte/in einer halben Stunde/bei/einer Erheberung/das Problem…</td>
</tr>
<tr>
<td></td>
<td>mism./SAVO</td>
<td>Der Psychologe/erkannte/eine halbe Stunde lang/bei/einer Erheberung/das Problem…</td>
</tr>
<tr>
<td>6 (4a) contr./SAVO</td>
<td>Der Kommissar/überführte/den Bankräuber/vor einer Stunde/beim/Banküberfall.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>coerc./SAVO</td>
<td>Der Kommissar/überführte/den Bankräuber/in einer Stunde/beim/Banküberfall.</td>
</tr>
<tr>
<td></td>
<td>mism./SAVO</td>
<td>Der Kommissar/überführte/den Bankräuber/eine Stunde lang/beim/Banküberfall.</td>
</tr>
<tr>
<td>6 (4b) contr./SAVO</td>
<td>Der Kommissar/überführte/vor einer Stunde/beim/Banküberfall/den Bankräuber…</td>
<td></td>
</tr>
<tr>
<td></td>
<td>coerc./SAVO</td>
<td>Der Kommissar/überführte/in einer Stunde/beim/Banküberfall/den Bankräuber…</td>
</tr>
<tr>
<td></td>
<td>mism./SAVO</td>
<td>Der Kommissar/überführte/eine Stunde lang/beim/Banküberfall/den Bankräuber…</td>
</tr>
<tr>
<td>item (Exp.)</td>
<td>condition</td>
<td>sentence</td>
</tr>
<tr>
<td>------------</td>
<td>-----------</td>
<td>---------</td>
</tr>
<tr>
<td>10 (4b)</td>
<td>contr./SVOA</td>
<td>Der Metzger/tötete/vor einer halben Stunde/im/städtischen Schlachthof/das Schwein…</td>
</tr>
<tr>
<td></td>
<td>coerc./SVOA</td>
<td>Der Metzger/tötete/in einer halben Stunde/im/städtischen Schlachthof/das Schwein…</td>
</tr>
<tr>
<td></td>
<td>mism./SVOA</td>
<td>Der Metzger/tötete/eine halbe Stunde lang/im/städtischen Schlachthof/das Schwein…</td>
</tr>
<tr>
<td>11 (4a)</td>
<td>contr./SVOA</td>
<td>Der Diplomand/erhielt/das Zeugnis/vor dreißig Minuten/auf/dem Dekanat.</td>
</tr>
<tr>
<td></td>
<td>coerc./SVOA</td>
<td>Der Diplomand/erhielt/das Zeugnis/in dreißig Minuten/auf/dem Dekanat.</td>
</tr>
<tr>
<td></td>
<td>mism./SVOA</td>
<td>Der Diplomand/erhielt/das Zeugnis/dreißig Minuten lang/auf/dem Dekanat.</td>
</tr>
<tr>
<td>11 (4b)</td>
<td>contr./SVOA</td>
<td>Der Diplomand/erhielt/vor dreißig Minuten/auf/dem Dekanat/das Zeugnis…</td>
</tr>
<tr>
<td></td>
<td>coerc./SVOA</td>
<td>Der Diplomand/erhielt/in dreißig Minuten/auf/dem Dekanat/das Zeugnis…</td>
</tr>
<tr>
<td></td>
<td>mism./SVOA</td>
<td>Der Diplomand/erhielt/dreißig Minuten lang/auf/dem Dekanat/das Zeugnis…</td>
</tr>
<tr>
<td>12 (4a)</td>
<td>contr./SVOA</td>
<td>Der Opa/fand/die Brille/vor fünf Minuten/in/einer Schublade.</td>
</tr>
<tr>
<td></td>
<td>coerc./SVOA</td>
<td>Der Opa/fand/die Brille/in fünf Minuten/in/einer Schublade.</td>
</tr>
<tr>
<td></td>
<td>mism./SVOA</td>
<td>Der Opa/fand/die Brille/fünf Minuten lang/in/einer Schublade.</td>
</tr>
<tr>
<td>12 (4b)</td>
<td>contr./SVOA</td>
<td>Der Opa/fand/vor fünf Minuten/in/einer Schublade/die Brille…</td>
</tr>
<tr>
<td></td>
<td>coerc./SVOA</td>
<td>Der Opa/fand/in fünf Minuten/in/einer Schublade/die Brille…</td>
</tr>
<tr>
<td></td>
<td>mism./SVOA</td>
<td>Der Opa/fand/fünf Minuten lang/in/einer Schublade/die Brille…</td>
</tr>
<tr>
<td>13 (4a)</td>
<td>contr./SVOA</td>
<td>Der Student/verstand/die Theorie/vor zwei Tagen/trotz/mangelnder Voraussetzungen.</td>
</tr>
<tr>
<td></td>
<td>coerc./SVOA</td>
<td>Der Student/verstand/die Theorie/in zwei Tagen/trotz/mangelnder Voraussetzungen.</td>
</tr>
<tr>
<td></td>
<td>mism./SVOA</td>
<td>Der Student/verstand/die Theorie/zwei Tage lang/trotz/mangelnder Voraussetzungen.</td>
</tr>
<tr>
<td>13 (4b)</td>
<td>contr./SVOA</td>
<td>Der Student/verstand/vor zwei Tagen/trotz/mangelnder Voraussetzungen/die Theorie…</td>
</tr>
<tr>
<td></td>
<td>coerc./SVOA</td>
<td>Der Student/verstand/in zwei Tagen/trotz/mangelnder Voraussetzungen/die Theorie…</td>
</tr>
<tr>
<td></td>
<td>mism./SVOA</td>
<td>Der Student/verstand/zwei Tage lang/trotz/mangelnder Voraussetzungen/die Theorie…</td>
</tr>
</tbody>
</table>
### Appendix G. Target sentences in Experiment 4 (4a and 4b)

<table>
<thead>
<tr>
<th>item (Exp.)</th>
<th>condition</th>
<th>sentence</th>
</tr>
</thead>
<tbody>
<tr>
<td>17 (4b)</td>
<td>contr./SAVO</td>
<td>Der Klemmpner/bemerkte/vor dreißig Minuten/bei/einer Inspektion/den Wasserrohrbruch…</td>
</tr>
<tr>
<td></td>
<td>coerc./SAVO</td>
<td>Der Klemmpner/bemerkte/in dreißig Minuten/bei/einer Inspektion/den Wasserrohrbruch…</td>
</tr>
<tr>
<td></td>
<td>mism./SAVO</td>
<td>Der Klemmpner/bemerkte/dreißig Minuten lang/bei/einer Inspektion/den Wasserrohrbruch…</td>
</tr>
<tr>
<td>18 (4a)</td>
<td>contr./SAVO</td>
<td>Der Höhlenforscher/verließ/die Tropfsteinhöhle/vor fünfzehn Minuten/durch/einen Kriechgang.</td>
</tr>
<tr>
<td></td>
<td>coerc./SAVO</td>
<td>Der Höhlenforscher/verließ/die Tropfsteinhöhle/in fünfzehn Minuten/durch/einen Kriechgang.</td>
</tr>
<tr>
<td></td>
<td>mism./SAVO</td>
<td>Der Höhlenforscher/verließ/die Tropfsteinhöhle/fünfzehn Minuten lang/durch/einen Kriechgang.</td>
</tr>
<tr>
<td>18 (4b)</td>
<td>contr./SAVO</td>
<td>Der Höhlenforscher/verließ/vor fünfzehn Minuten/durch/einen Kriechgang/die Tropfsteinhöhle…</td>
</tr>
<tr>
<td></td>
<td>coerc./SAVO</td>
<td>Der Höhlenforscher/verließ/in fünfzehn Minuten/durch/einen Kriechgang/die Tropfsteinhöhle…</td>
</tr>
<tr>
<td></td>
<td>mism./SAVO</td>
<td>Der Höhlenforscher/verließ/fünfzehn Minuten lang/durch/einen Kriechgang/die Tropfsteinhöhle…</td>
</tr>
<tr>
<td>19 (4a)</td>
<td>contr./SAVO</td>
<td>Der Schüler/bewältigte/die Aufgabe/vor einer dreiviertel Stunde/bei/der Mathearbeit.</td>
</tr>
<tr>
<td></td>
<td>coerc./SAVO</td>
<td>Der Schüler/bewältigte/die Aufgabe/in einer dreiviertel Stunde/bei/der Mathearbeit.</td>
</tr>
<tr>
<td></td>
<td>mism./SAVO</td>
<td>Der Schüler/bewältigte/die Aufgabe/eine dreiviertel Stunde lang/bei/der Mathearbeit.</td>
</tr>
<tr>
<td>19 (4b)</td>
<td>contr./SAVO</td>
<td>Der Schüler/bewältigte/vor einer dreiviertel Stunde/bei/der Mathearbeit/die Aufgabe…</td>
</tr>
<tr>
<td></td>
<td>coerc./SAVO</td>
<td>Der Schüler/bewältigte/in einer dreiviertel Stunde/bei/der Mathearbeit/die Aufgabe…</td>
</tr>
<tr>
<td></td>
<td>mism./SAVO</td>
<td>Der Schüler/bewältigte/eine dreiviertel Stunde lang/bei/der Mathearbeit/die Aufgabe…</td>
</tr>
<tr>
<td>20 (4a)</td>
<td>contr./SAVO</td>
<td>Der Mathematiker/durchschaute/den Beweis/vor zwei Tagen/trotz/anfänglicher Verständnisprobleme.</td>
</tr>
<tr>
<td></td>
<td>coerc./SAVO</td>
<td>Der Mathematiker/durchschaute/den Beweis/in zwei Tagen/trotz/anfänglicher Verständnisprobleme.</td>
</tr>
<tr>
<td></td>
<td>mism./SAVO</td>
<td>Der Mathematiker/durchschaute/den Beweis/zu zwei Tagen lang/trotz/anfänglicher Verständnisprobleme.</td>
</tr>
<tr>
<td>20 (4b)</td>
<td>contr./SAVO</td>
<td>Der Mathematiker/durchschaute/vor zwei Tagen/trotz/anfänglicher Verständnisprobleme/den Beweis…</td>
</tr>
<tr>
<td></td>
<td>coerc./SAVO</td>
<td>Der Mathematiker/durchschaute/in zwei Tagen/trotz/anfänglicher Verständnisprobleme/den Beweis…</td>
</tr>
<tr>
<td></td>
<td>mism./SAVO</td>
<td>Der Mathematiker/durchschaute/zwei Tage lang/trotz/anfänglicher Verständnisprobleme/den Beweis…</td>
</tr>
<tr>
<td>21 (4a)</td>
<td>contr./SAVO</td>
<td>Die NASA/startete/die Rakete/vor einer Stunde/trotz/schlechten Wetters.</td>
</tr>
<tr>
<td></td>
<td>coerc./SAVO</td>
<td>Die NASA/startete/die Rakete/in einer Stunde/trotz/schlechten Wetters.</td>
</tr>
<tr>
<td></td>
<td>mism./SAVO</td>
<td>Die NASA/startete/die Rakete/eine Stunde lang/trotz/schlechten Wetters.</td>
</tr>
<tr>
<td>21 (4b)</td>
<td>contr./SAVO</td>
<td>Die NASA/startete/vor einer Stunde/trotz/schlechten Wetters/die Rakete…</td>
</tr>
<tr>
<td></td>
<td>coerc./SAVO</td>
<td>Die NASA/startete/in einer Stunde/trotz/schlechten Wetters/die Rakete…</td>
</tr>
<tr>
<td></td>
<td>mism./SAVO</td>
<td>Die NASA/startete/eine Stunde lang/trotz/schlechten Wetters/die Rakete…</td>
</tr>
<tr>
<td>22 (4a)</td>
<td>contr./SAVO</td>
<td>Die Grippe/befiel/den Patienten/vor vierundzwanzig Stunden/trotz/Quarantänaufnahmen.</td>
</tr>
<tr>
<td></td>
<td>coerc./SAVO</td>
<td>Die Grippe/befiel/den Patienten/in vierundzwanzig Stunden/trotz/Quarantänaufnahmen.</td>
</tr>
<tr>
<td></td>
<td>mism./SAVO</td>
<td>Die Grippe/befiel/vierundzwanzig Stunden lang/trotz/Quarantänaufnahmen/den Patienten…</td>
</tr>
<tr>
<td>22 (4b)</td>
<td>contr./SAVO</td>
<td>Die Grippe/befiel/vor vierundzwanzig Stunden/trotz/Quarantänaufnahmen/den Patienten…</td>
</tr>
<tr>
<td></td>
<td>coerc./SAVO</td>
<td>Die Grippe/befiel/in vierundzwanzig Stunden/trotz/Quarantänaufnahmen/den Patienten…</td>
</tr>
<tr>
<td></td>
<td>mism./SAVO</td>
<td>Die Grippe/befiel/vierundzwanzig Stunden lang/trotz/Quarantänaufnahmen/den Patienten…</td>
</tr>
<tr>
<td>23 (4a)</td>
<td>contr./SAVO</td>
<td>Der Ringer/übermannte/den Gegner/vor einer halben Stunde/bei/einem Schaukampf.</td>
</tr>
<tr>
<td></td>
<td>coerc./SAVO</td>
<td>Der Ringer/übermannte/den Gegner/in einer halben Stunde/bei/einem Schaukampf.</td>
</tr>
<tr>
<td></td>
<td>mism./SAVO</td>
<td>Der Ringer/übermannte/den Gegner/eine halbe Stunde lang/bei/einem Schaukampf.</td>
</tr>
<tr>
<td>23 (4b)</td>
<td>contr./SAVO</td>
<td>Der Ringer/übermannte/vor einer halben Stunde/bei/ einem Schaukampf/den Gegner…</td>
</tr>
<tr>
<td></td>
<td>coerc./SAVO</td>
<td>Der Ringer/übermannte/in einer halben Stunde/bei/ einem Schaukampf/den Gegner…</td>
</tr>
<tr>
<td></td>
<td>mism./SAVO</td>
<td>Der Ringer/übermannte/eine halbe Stunde lang/bei/ einem Schaukampf/den Gegner…</td>
</tr>
<tr>
<td>24 (4a)</td>
<td>contr./SAVO</td>
<td>Der Rennfahrer/besiegte/die Konkurrenz/vor dreißig Minuten/auf/dem Nürburgring.</td>
</tr>
<tr>
<td></td>
<td>coerc./SAVO</td>
<td>Der Rennfahrer/besiegte/die Konkurrenz/in dreißig Minuten/auf/dem Nürburgring.</td>
</tr>
<tr>
<td></td>
<td>mism./SAVO</td>
<td>Der Rennfahrer/besiegte/die Konkurrenz/dreißig Minuten lang/auf/dem Nürburgring.</td>
</tr>
<tr>
<td>item (Exp.)</td>
<td>condition</td>
<td>sentence</td>
</tr>
<tr>
<td>------------</td>
<td>-----------</td>
<td>----------</td>
</tr>
<tr>
<td>24 (4b)</td>
<td>contr./SVOA</td>
<td>Der Rennfahrer/besiegte/vor dreißig Minuten/auf/dem Nürburgring/die Konkurrenz…</td>
</tr>
<tr>
<td></td>
<td>coerc./SVOA</td>
<td>Der Rennfahrer/besiegte/in dreißig Minuten/auf/dem Nürburgring/die Konkurrenz…</td>
</tr>
<tr>
<td></td>
<td>mism./SVOA</td>
<td>Der Rennfahrer/besiegte/dreißig Minuten lang/auf/dem Nürburgring/die Konkurrenz…</td>
</tr>
<tr>
<td>25 (4a)</td>
<td>contr./SVOA</td>
<td>Der Matrose/erspähte/die Insel/vor fünf Minuten/von/der Spitze des Masts.</td>
</tr>
<tr>
<td></td>
<td>coerc./SVOA</td>
<td>Der Matrose/erspähte/die Insel/in fünf Minuten/von/der Spitze des Masts.</td>
</tr>
<tr>
<td></td>
<td>mism./SVOA</td>
<td>Der Matrose/erspähte/die Insel/fünf Minuten lang/von/der Spitze des Masts.</td>
</tr>
<tr>
<td>25 (4b)</td>
<td>contr./SVOA</td>
<td>Der Matrose/erspähte/vor fünf Minuten/von/der Spitze des Masts/die Insel…</td>
</tr>
<tr>
<td></td>
<td>coerc./SVOA</td>
<td>Der Matrose/erspähte/in fünf Minuten/von/der Spitze des Masts/die Insel…</td>
</tr>
<tr>
<td></td>
<td>mism./SVOA</td>
<td>Der Matrose/erspähte/fünf Minuten lang/von/der Spitze des Masts/die Insel…</td>
</tr>
<tr>
<td>26 (4a)</td>
<td>contr./SVOA</td>
<td>Der Konferenzzugehöriger/erblichte/seinen Kollegen/vor zehn Minuten/im/Publikum.</td>
</tr>
<tr>
<td></td>
<td>coerc./SVOA</td>
<td>Der Konferenzzugehöriger/erblichte/seinen Kollegen/in zehn Minuten/im/Publikum.</td>
</tr>
<tr>
<td></td>
<td>mism./SVOA</td>
<td>Der Konferenzzugehöriger/erblichte/seinen Kollegen/zehn Minuten lang/im/Publikum.</td>
</tr>
<tr>
<td>26 (4b)</td>
<td>contr./SVOA</td>
<td>Der Konferenzzugehöriger/erblichte/vor zehn Minuten/im/Publikum/seinen Kollegen…</td>
</tr>
<tr>
<td></td>
<td>coerc./SVOA</td>
<td>Der Konferenzzugehöriger/erblichte/in zehn Minuten/im/Publikum/seinen Kollegen…</td>
</tr>
<tr>
<td></td>
<td>mism./SVOA</td>
<td>Der Konferenzzugehöriger/erblichte/zehn Minuten lang/im/Publikum/seinen Kollegen…</td>
</tr>
<tr>
<td>27 (4a)</td>
<td>contr./SVOA</td>
<td>Die Schülerin/verblüffte/ihrer Lehrer/vor zwanzig Minuten/mit/der korrekten Lösung.</td>
</tr>
<tr>
<td></td>
<td>coerc./SVOA</td>
<td>Die Schülerin/verblüffte/ihrer Lehrer/in zwanzig Minuten/mit/der korrekten Lösung.</td>
</tr>
<tr>
<td></td>
<td>mism./SVOA</td>
<td>Die Schülerin/verblüffte/ihrer Lehrer/zwanzig Minuten lang/mit/der korrekten Lösung.</td>
</tr>
<tr>
<td>27 (4b)</td>
<td>contr./SVOA</td>
<td>Die Schülerin/verblüffte/vor zwanzig Minuten/mit/der korrekten Lösung/ihrer Lehrer…</td>
</tr>
<tr>
<td></td>
<td>coerc./SVOA</td>
<td>Die Schülerin/verblüffte/in zwanzig Minuten/mit/der korrekten Lösung/ihrer Lehrer…</td>
</tr>
<tr>
<td></td>
<td>mism./SVOA</td>
<td>Die Schülerin/verblüffte/zwanzig Minuten lang/mit/der korrekten Lösung/ihrer Lehrer…</td>
</tr>
</tbody>
</table>
### Target sentences in Experiment 5

<table>
<thead>
<tr>
<th>Item</th>
<th>Condition</th>
<th>Sentence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>svoa-for</td>
<td>Johann/errichtete/das Haus/zwei Jahre lang/trotz/finanzieller Probleme.</td>
</tr>
<tr>
<td></td>
<td>svoa-in</td>
<td>Johann/errichtete/das Haus/in zwei Jahren/trotz/finanzieller Probleme.</td>
</tr>
<tr>
<td>2</td>
<td>svoa-for</td>
<td>Der Autor/verfasste/den Roman/zwei Jahre lang/trotz/starker Kritik an dem Projekt.</td>
</tr>
<tr>
<td></td>
<td>svoa-in</td>
<td>Der Autor/verfasste/den Roman/in zwei Jahren/trotz/starker Kritik an dem Projekt.</td>
</tr>
<tr>
<td>3</td>
<td>svoa-for</td>
<td>Der Restauranbesucher/verspeiste/das Menü/zwanzig Minuten lang/in/großer Eile.</td>
</tr>
<tr>
<td></td>
<td>svoa-in</td>
<td>Der Restauranbesucher/verspeiste/das Menü/in zwanzig Minuten/in/großer Eile.</td>
</tr>
<tr>
<td></td>
<td>svao-in</td>
<td>Der Restauranbesucher/verspeiste/in zwanzig Minuten/in/großer Eile/das Menü/von/der Speisekarte.</td>
</tr>
<tr>
<td>4</td>
<td>svoa-for</td>
<td>Der Hund/verschlang/das Futter/fünf Minuten lang/mit/großem Appetit.</td>
</tr>
<tr>
<td></td>
<td>svoa-in</td>
<td>Der Hund/verschlang/das Futter/in fünf Minuten/mit/großem Appetit.</td>
</tr>
<tr>
<td></td>
<td>svao-for</td>
<td>Der Hund/verschlang/fünf Minuten lang/mit/großem Appetit/das Futter/aus/dem Fressnapf.</td>
</tr>
<tr>
<td></td>
<td>svao-in</td>
<td>Der Hund/verschlang/in fünf Minuten/mit/großem Appetit/das Futter/aus/dem Fressnapf.</td>
</tr>
</tbody>
</table>
### Appendix H. Target sentences in Experiment 5

<table>
<thead>
<tr>
<th>item</th>
<th>condition</th>
<th>sentence</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>svoa-for</td>
<td>Der Informatiker entschlüsselte den Code zwanzig Minuten lang mit großem Eifer.</td>
</tr>
<tr>
<td></td>
<td>svoa-in</td>
<td>Der Informatiker entschlüsselte den Code in zwanzig Minuten mit großem Eifer.</td>
</tr>
<tr>
<td>6</td>
<td>svoa-for</td>
<td>Der Mann entlud den LKW zwei Stunden lang trotz schlechten Wetters mit Anhänger.</td>
</tr>
<tr>
<td></td>
<td>svoa-in</td>
<td>Der Mann entlud den LKW in zwei Stunden trotz schlechten Wetters mit Anhänger.</td>
</tr>
<tr>
<td>7</td>
<td>svoa-for</td>
<td>Der Kommissar überführte den Dieb eine Woche lang beim Verhör aus Sizilien.</td>
</tr>
<tr>
<td></td>
<td>svoa-in</td>
<td>Der Kommissar überführte den Dieb in einer Woche beim Verhör aus Sizilien.</td>
</tr>
<tr>
<td>8</td>
<td>svoa-for</td>
<td>Der Athlet absolvierte den Lauf zwanzig Minuten lang bei der letzten Olympiade mit spaltenreichen Gletscher.</td>
</tr>
<tr>
<td></td>
<td>svoa-in</td>
<td>Der Athlet absolvierte den Lauf in zwanzig Minuten bei der letzten Olympiade mit spaltenreichen Gletscher.</td>
</tr>
<tr>
<td>9</td>
<td>svoa-for</td>
<td>Der Architekt erstellte den Plan einer Woche lang an seinem CAD Arbeitsplatz.</td>
</tr>
<tr>
<td></td>
<td>svoa-in</td>
<td>Der Architekt erstellte den Plan in einer Woche an seinem CAD Arbeitsplatz.</td>
</tr>
<tr>
<td>10</td>
<td>svoa-for</td>
<td>Die Armee zerstörte die Stadt sieben Tage lang während des dreißigjährigen Kriegs.</td>
</tr>
<tr>
<td></td>
<td>svoa-in</td>
<td>Die Armee zerstörte die Stadt in sieben Tagen während des dreißigjährigen Kriegs.</td>
</tr>
</tbody>
</table>
APPENDIX I

Target sentences in Experiment 6

<table>
<thead>
<tr>
<th>item</th>
<th>condition</th>
<th>sentence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>sva-for</td>
<td>Das Licht/flackerte/den ganzen Abend/am vergangenen Dienstag/in der Nebenwohnung, …</td>
</tr>
<tr>
<td></td>
<td>sva-ago</td>
<td>Das Licht/flackerte/am späten Abend/am vergangenen Dienstag/in der Nebenwohnung, …</td>
</tr>
<tr>
<td></td>
<td>vas-for</td>
<td>In der Nebenwohnung/flackerte/den ganzen Abend/am vergangenen Dienstag/das Licht, …</td>
</tr>
<tr>
<td></td>
<td>vas-ago</td>
<td>In der Nebenwohnung/flackerte/am späten Abend/am vergangenen Dienstag/das Licht, …</td>
</tr>
<tr>
<td>2</td>
<td>sva-for</td>
<td>Die Kanone/donnerte/den ganzen Morgen/beim Faschingsumzug/am Rosenmontag, …</td>
</tr>
<tr>
<td></td>
<td>sva-ago</td>
<td>Die Kanone/donnerte/am frühen Morgen/beim Faschingsumzug/am Rosenmontag, …</td>
</tr>
<tr>
<td></td>
<td>vas-for</td>
<td>Am Rosenmontag/donnerte/den ganzen Morgen/beim Faschingsumzug/die Kanone, …</td>
</tr>
<tr>
<td></td>
<td>vas-ago</td>
<td>Am Rosenmontag/donnerte/am frühen Morgen/beim Faschingsumzug/die Kanone, …</td>
</tr>
<tr>
<td>3</td>
<td>sva-for</td>
<td>Die Trommel/ertönte/tagelang/bei einem Wochenendkurs/in der Volkshochschule, …</td>
</tr>
<tr>
<td></td>
<td>sva-ago</td>
<td>Die Trommel/ertönte/neulich/bei einem Wochenendkurs/in der Volkshochschule, …</td>
</tr>
<tr>
<td></td>
<td>vas-for</td>
<td>In der Volkshochschule/ertönte/tagelang/bei einem Wochenendkurs/die Trommel, …</td>
</tr>
<tr>
<td></td>
<td>vas-ago</td>
<td>In der Volkshochschule/ertönte/neulich/bei einem Wochenendkurs/die Trommel, …</td>
</tr>
<tr>
<td>4</td>
<td>sva-for</td>
<td>Das Kriegsgeheul/erscholl/wochenlang/bei den Yanomami-Indianern/im Amazonasgebiet, …</td>
</tr>
<tr>
<td></td>
<td>sva-ago</td>
<td>Das Kriegsgeheul/erscholl/letzte Woche/bei den Yanomami-Indianern/im Amazonasgebiet, …</td>
</tr>
<tr>
<td></td>
<td>vas-for</td>
<td>Im Amazonasgebiet/erscholl/wochenlang/bei den Yanomami-Indianern/das Kriegsgeheul, …</td>
</tr>
<tr>
<td></td>
<td>vas-ago</td>
<td>Im Amazonasgebiet/erscholl/letzte Woche/bei den Yanomami-Indianern/das Kriegsgeheul, …</td>
</tr>
</tbody>
</table>
The Processing of Events

<table>
<thead>
<tr>
<th>item</th>
<th>condition</th>
<th>sentence</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>sva-for</td>
<td>Der Wecker/klingelte/den ganzen Dienstag/in einem Zimmer/im Studentenwohnheim, …</td>
</tr>
<tr>
<td></td>
<td>sva-ago</td>
<td>Der Wecker/klingelte/um drei Uhr morgens/in einem Zimmer/im Studentenwohnheim, …</td>
</tr>
<tr>
<td></td>
<td>vas-for</td>
<td>Im Studentenwohnheim/klingelte/den ganzen Dienstag/in einem Zimmer/der Wecker, …</td>
</tr>
<tr>
<td></td>
<td>vas-ago</td>
<td>Im Studentenwohnheim/klingelte/um drei Uhr morgens/in einem Zimmer/der Wecker, …</td>
</tr>
<tr>
<td>6</td>
<td>sva-for</td>
<td>Eine Studentin/nieste/mehrere Tage/überaus laut/auf der Exkursion, …</td>
</tr>
<tr>
<td></td>
<td>sva-ago</td>
<td>Eine Studentin/nieste/gerade eben/überaus laut/auf der Exkursion, …</td>
</tr>
<tr>
<td></td>
<td>vas-for</td>
<td>Auf der Exkursion/nieste/mehrere Tage/überaus laut/eine Studentin, …</td>
</tr>
<tr>
<td></td>
<td>vas-ago</td>
<td>Auf der Exkursion/nieste/gerade eben/überaus laut/eine Studentin, …</td>
</tr>
<tr>
<td>7</td>
<td>sva-for</td>
<td>Die Alarmglocke/schrillte/die ganze Nacht/in den letzten Kriegstagen/im zweiten Weltkrieg, …</td>
</tr>
<tr>
<td></td>
<td>sva-ago</td>
<td>Die Alarmglocke/schrillte/in einer Mainacht/in den letzten Kriegstagen/der Alarmglocke, …</td>
</tr>
<tr>
<td></td>
<td>vas-for</td>
<td>Im zweiten Weltkrieg/schrillte/die ganze Nacht/in den letzten Kriegstagen/der Alarmglocke, …</td>
</tr>
<tr>
<td></td>
<td>vas-ago</td>
<td>Im zweiten Weltkrieg/schrillte/in einer Mainacht/in den letzten Kriegstagen/der Alarmglocke, …</td>
</tr>
<tr>
<td>8</td>
<td>sva-for</td>
<td>Der Telegraph/klackte/fünf Minuten lang/in einer entlegenen Station/im amerikanischen Westen, …</td>
</tr>
<tr>
<td></td>
<td>sva-ago</td>
<td>Der Telegraph/klackte/laut vernehmlich/in einer entlegenen Station/im amerikanischen Westen, …</td>
</tr>
<tr>
<td></td>
<td>vas-for</td>
<td>Im amerikanischen Westen/klackte/fünf Minuten lang/in einer entlegenen Station/der Telegraph, …</td>
</tr>
<tr>
<td></td>
<td>vas-ago</td>
<td>Im amerikanischen Westen/klackte/laut vernehmlich/in einer entlegenen Station/der Telegraph, …</td>
</tr>
<tr>
<td>9</td>
<td>sva-for</td>
<td>Das Bett/knarrte/mehrere Jahre lang/unangenehm laut/im Haus meines Bruders, …</td>
</tr>
<tr>
<td></td>
<td>sva-ago</td>
<td>Das Bett/knarrte/vorgestern früh/unangenehm laut/im Haus meines Bruders, …</td>
</tr>
<tr>
<td></td>
<td>vas-for</td>
<td>Im Haus meines Bruders/knarrte/mehrere Jahre lang/unangenehm laut/das Bett, …</td>
</tr>
<tr>
<td></td>
<td>vas-ago</td>
<td>Im Haus meines Bruders/knarrte/vorgestern früh/unangenehm laut/das Bett, …</td>
</tr>
<tr>
<td>10</td>
<td>sva-for</td>
<td>Das Gewehr/knallte/die ganze Nacht/während einer Schlacht/im ersten Weltkrieg, …</td>
</tr>
<tr>
<td></td>
<td>sva-ago</td>
<td>Das Gewehr/knallte/mitten in der Nacht/während einer Schlacht/im ersten Weltkrieg, …</td>
</tr>
<tr>
<td></td>
<td>vas-for</td>
<td>Im ersten Weltkrieg/knallte/die ganze Nacht/während einer Schlacht/das Gewehr, …</td>
</tr>
<tr>
<td></td>
<td>vas-ago</td>
<td>Im ersten Weltkrieg/knallte/mitten in der Nacht/während einer Schlacht/das Gewehr, …</td>
</tr>
</tbody>
</table>

Appendix I. Target sentences in Experiment 6

<table>
<thead>
<tr>
<th>item</th>
<th>condition</th>
<th>sentence</th>
</tr>
</thead>
<tbody>
<tr>
<td>11</td>
<td>sva-for</td>
<td>Die Messnadel/zuckte/wochenlang/in einer seismographischen Station/auf Reunion, …</td>
</tr>
<tr>
<td></td>
<td>sva-ago</td>
<td>Die Messnadel/zuckte/um zehn Uhr/in einer seismographischen Station/auf Reunion, …</td>
</tr>
<tr>
<td></td>
<td>vas-for</td>
<td>Auf Reunion/zuckte/wochenlang/in einer seismographischen Station/die Messnadel, …</td>
</tr>
<tr>
<td></td>
<td>vas-ago</td>
<td>Auf Reunion/zuckte/um zehn Uhr/in einer seismographischen Station/die Messnadel, …</td>
</tr>
<tr>
<td>12</td>
<td>sva-for</td>
<td>Der Überwachungsmonitor/piepste/wochenlang/in Zimmer fünf/in der Chirurgie, …</td>
</tr>
<tr>
<td></td>
<td>sva-ago</td>
<td>Der Überwachungsmonitor/piepste/vorgestern/in Zimmer fünf/in der Chirurgie, …</td>
</tr>
<tr>
<td></td>
<td>vas-for</td>
<td>In der Chirurgie/piepste/wochenlang/in Zimmer fünf/der Überwachungsmonitor, …</td>
</tr>
<tr>
<td></td>
<td>vas-ago</td>
<td>In der Chirurgie/piepste/vorgestern/in Zimmer fünf/der Überwachungsmonitor, …</td>
</tr>
<tr>
<td>13</td>
<td>sva-for</td>
<td>Ein Hochhaus/wackelte/eine Stunde lang/in einem Geschäftsviertel/in Los Angeles, …</td>
</tr>
<tr>
<td></td>
<td>sva-ago</td>
<td>Ein Hochhaus/wackelte/heute Morgen/in einem Geschäftsviertel/in Los Angeles, …</td>
</tr>
<tr>
<td></td>
<td>vas-for</td>
<td>In Los Angeles/wackelte/eine Stunde lang/in einem Geschäftsviertel/ein Hochhaus, …</td>
</tr>
<tr>
<td></td>
<td>vas-ago</td>
<td>In Los Angeles/wackelte/heute Morgen/in einem Geschäftsviertel/ein Hochhaus, …</td>
</tr>
<tr>
<td>14</td>
<td>sva-for</td>
<td>Die Quelle/gluckerte/mehrere Tage lang/nach schweren Regenfällen/laut und vernehmlich, …</td>
</tr>
<tr>
<td></td>
<td>sva-ago</td>
<td>Die Quelle/gluckerte/am Montagmorgen/nach schweren Regenfällen/laut und vernehmlich, …</td>
</tr>
<tr>
<td></td>
<td>vas-for</td>
<td>Laut und vernehmlich/gluckerte/mehrere Tage lang/nach schweren Regenfällen/die Quelle, …</td>
</tr>
<tr>
<td></td>
<td>vas-ago</td>
<td>Laut und vernehmlich/gluckerte/am Montagmorgen/nach schweren Regenfällen/die Quelle, …</td>
</tr>
<tr>
<td>15</td>
<td>sva-for</td>
<td>Das Wasser/spritzte/den ganzen Tag/im Pool/im Nachbargarten, …</td>
</tr>
<tr>
<td></td>
<td>sva-ago</td>
<td>Das Wasser/spritzte/so gegen vier/im Pool/im Nachbargarten, …</td>
</tr>
<tr>
<td></td>
<td>vas-for</td>
<td>Im Nachbargarten/spritzte/den ganzen Tag/im Pool/das Wasser, …</td>
</tr>
<tr>
<td></td>
<td>vas-ago</td>
<td>Im Nachbargarten/spritzte/so gegen vier/im Pool/das Wasser, …</td>
</tr>
<tr>
<td>16</td>
<td>sva-for</td>
<td>Eine Sirene/schirllte/das ganze Spiel über/ohrenbetäubend laut/im Fanblock, …</td>
</tr>
<tr>
<td></td>
<td>sva-ago</td>
<td>Eine Sirene/schirllte/gestern Abend/ohrenbetäubend laut/im Fanblock, …</td>
</tr>
<tr>
<td></td>
<td>vas-for</td>
<td>Im Fanblock/schirllte/das ganze Spiel über/ohrenbetäubend laut/eine Sirene, …</td>
</tr>
<tr>
<td></td>
<td>vas-ago</td>
<td>Im Fanblock/schirllte/gestern Abend/ohrenbetäubend laut/eine Sirene, …</td>
</tr>
</tbody>
</table>
The Processing of Events

<table>
<thead>
<tr>
<th>item</th>
<th>condition</th>
<th>sentence</th>
</tr>
</thead>
<tbody>
<tr>
<td>17</td>
<td>sva-for</td>
<td>Das Siegesgeschrei/erdröhnte/die gesamte Partie/beim WM-Finale/im ausverkauften Stadion, …</td>
</tr>
<tr>
<td></td>
<td>sva-ago</td>
<td>Das Siegesgeschrei/erdröhnte/während der Partie/beim WM-Finale/im ausverkauften Stadion, …</td>
</tr>
<tr>
<td></td>
<td>vas-for</td>
<td>Im ausverkauften Stadion/erdröhnte/die gesamte Partie/beim WM-Finale/das Siegesgeschrei, …</td>
</tr>
<tr>
<td></td>
<td>vas-ago</td>
<td>Im ausverkauften Stadion/erdröhnte/während der Partie/beim WM-Finale/das Siegesgeschrei, …</td>
</tr>
<tr>
<td>18</td>
<td>sva-for</td>
<td>Die Ölwanne/tropfte/ein ganzes Jahr lang/bei Peters altem Ford/auf den Garagenboden, …</td>
</tr>
<tr>
<td></td>
<td>sva-ago</td>
<td>Die Ölwanne/tropfte/im vorigen Jahr/bei Peters altem Ford/auf den Garagenboden, …</td>
</tr>
<tr>
<td></td>
<td>vas-for</td>
<td>Auf den Garagenboden/tropfte/ein ganzes Jahr lang/bei Peters altem Ford/die Ölwanne, …</td>
</tr>
<tr>
<td></td>
<td>vas-ago</td>
<td>Auf den Garagenboden/tropfte/im vorigen Jahr/bei Peters altem Ford/die Ölwanne, …</td>
</tr>
<tr>
<td>19</td>
<td>sva-for</td>
<td>Der Schlachtruf/ertönte/tagelang/in einer Schlucht/in den Schweizer Bergen, …</td>
</tr>
<tr>
<td></td>
<td>sva-ago</td>
<td>Der Schlachtruf/ertönte/einstmals/in einer Schlucht/in den Schweizer Bergen, …</td>
</tr>
<tr>
<td></td>
<td>vas-for</td>
<td>In den Schweizer Bergen/ertönte/tagelang/in einer Schlucht/der Schlachtruf, …</td>
</tr>
<tr>
<td></td>
<td>vas-ago</td>
<td>In den Schweizer Bergen/ertönte/einstmals/in einer Schlucht/der Schlachtruf, …</td>
</tr>
<tr>
<td>20</td>
<td>sva-for</td>
<td>Der Gipfel/erglühte/fast den ganzen Sommer/rötlich/in der Abendsonne, …</td>
</tr>
<tr>
<td></td>
<td>sva-ago</td>
<td>Der Gipfel/erglühte/beim Sonnenuntergang/rötlich/in der Abendsonne, …</td>
</tr>
<tr>
<td></td>
<td>vas-for</td>
<td>In der Abendsonne/erglühte/fast den ganzen Sommer/rötlich/der Gipfel, …</td>
</tr>
<tr>
<td></td>
<td>vas-ago</td>
<td>In der Abendsonne/erglühte/beim Sonnenuntergang/rötlich/der Gipfel, …</td>
</tr>
<tr>
<td>21</td>
<td>sva-for</td>
<td>Der Wohnwagen/schlingerte/die ganze Fahrt über/in engen Kurven/beim Bergabwärtsfahren, …</td>
</tr>
<tr>
<td></td>
<td>sva-ago</td>
<td>Der Wohnwagen/schlingerte/letzte Woche/in einer engen Kurve/beim Bergabwärtsfahren, …</td>
</tr>
<tr>
<td></td>
<td>vas-for</td>
<td>Beim Bergabwärtsfahren/schlingerte/die ganze Fahrt über/in engen Kurven/der Wohnwagen, …</td>
</tr>
<tr>
<td></td>
<td>vas-ago</td>
<td>Beim Bergabwärtsfahren/schlingerte/letzte Woche/in einer engen Kurve/der Wohnwagen, …</td>
</tr>
<tr>
<td>22</td>
<td>sva-for</td>
<td>Der kleine Hans/hüpfte/den ganzen Mittag/mit Begeisterung/über das Seil, …</td>
</tr>
<tr>
<td></td>
<td>sva-ago</td>
<td>Der kleine Hans/hüpfte/heute Nachmittag/mit Begeisterung/über das Seil, …</td>
</tr>
</tbody>
</table>

Appendix I. Target sentences in Experiment 6

<table>
<thead>
<tr>
<th>item</th>
<th>condition</th>
<th>sentence</th>
</tr>
</thead>
<tbody>
<tr>
<td>22</td>
<td>vas-for</td>
<td>Über das Seil/hüpfte/den ganzen Mittag/mit Begeisterung/der kleine Hans, …</td>
</tr>
<tr>
<td></td>
<td>vas-ago</td>
<td>Über das Seil/hüpfte/heute Nachmittag/mit Begeisterung/der kleine Hans, …</td>
</tr>
<tr>
<td>23</td>
<td>sva-for</td>
<td>Unser Nachbar/rauchte/den ganzen Vormittag/auf der Terrasse/vor dem Haus, …</td>
</tr>
<tr>
<td></td>
<td>sva-ago</td>
<td>Unser Nachbar/rauchte/vor einer halben Stunde/auf der Terrasse/vor dem Haus, …</td>
</tr>
<tr>
<td></td>
<td>vas-for</td>
<td>Vor dem Haus/rauchte/den ganzen Vormittag/auf der Terrasse/unser Nachbar, …</td>
</tr>
<tr>
<td></td>
<td>vas-ago</td>
<td>Vor dem Haus/rauchte/vor einer halben Stunde/auf der Terrasse/unser Nachbar, …</td>
</tr>
<tr>
<td>24</td>
<td>sva-for</td>
<td>Der kleine Junge/gähnte/den ganzen Abend/vor dem Fernseher/im Wohnzimmer, …</td>
</tr>
<tr>
<td></td>
<td>sva-ago</td>
<td>Der kleine Junge/gähnte/vor wenigen Minuten/vor dem Fernseher/im Wohnzimmer, …</td>
</tr>
<tr>
<td></td>
<td>vas-for</td>
<td>Im Wohnzimmer/gähnte/den ganzen Abend/vor dem Fernseher/der kleine Junge, …</td>
</tr>
<tr>
<td></td>
<td>vas-ago</td>
<td>Im Wohnzimmer/gähnte/vor wenigen Minuten/vor dem Fernseher/der kleine Junge, …</td>
</tr>
</tbody>
</table>

note: sva = subj. > v. > adv.; vas = v. > adv. > subj.; for = for-adverbial; ago = x time ago
Appendix J

Target sentences in Experiment 8

Table J.1. Items used in Experiment 8

<table>
<thead>
<tr>
<th>item</th>
<th>condition</th>
<th>sentence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>mm-sg</td>
<td>Den ersten Preis/beim Rudern/gewann/knapp drei Stunden lang/eine kanadische/Athletin.</td>
</tr>
<tr>
<td></td>
<td>cc-sg</td>
<td>Den ersten Preis/beim Rudern/gewann/in knapp drei Stunden/eine kanadische/Athletin.</td>
</tr>
<tr>
<td></td>
<td>ctrl-sg</td>
<td>Den ersten Preis/beim Rudern/gewann/vor knapp drei Stunden/eine kanadische/Athletin.</td>
</tr>
<tr>
<td></td>
<td>mm-pl</td>
<td>Den ersten Preis/beim Rudern/gewannen/knapp drei Stunden lang/die kanadischen/Athleten.</td>
</tr>
<tr>
<td></td>
<td>cc-pl</td>
<td>Den ersten Preis/beim Rudern/gewannen/in knapp drei Stunden/die kanadischen/Athleten.</td>
</tr>
<tr>
<td></td>
<td>ctrl-pl</td>
<td>Den ersten Preis/beim Rudern/gewannen/vor knapp drei Stunden/die kanadischen/Athleten.</td>
</tr>
<tr>
<td>2</td>
<td>mm-sg</td>
<td>Den Hafen/von San Francisco/erreichte/eine halbe Stunde lang/eine japanische/Luxusyacht.</td>
</tr>
<tr>
<td></td>
<td>cc-sg</td>
<td>Den Hafen/von San Francisco/erreichte/in einer halben Stunde/eine japanische/Luxusyacht.</td>
</tr>
<tr>
<td></td>
<td>ctrl-sg</td>
<td>Den Hafen/von San Francisco/erreichte/vor einer halben Stunde/eine japanische/Luxusyacht.</td>
</tr>
<tr>
<td></td>
<td>mm-pl</td>
<td>Den Hafen/von San Francisco/erreichten/eine halbe Stunde lang/mehrere japanische/Luxusyachten.</td>
</tr>
<tr>
<td></td>
<td>cc-pl</td>
<td>Den Hafen/von San Francisco/erreichten/in einer halben Stunde/mehrere japanische/Luxusyachten.</td>
</tr>
<tr>
<td></td>
<td>ctrl-pl</td>
<td>Den Hafen/von San Francisco/erreichten/vor einer halben Stunde/mehrere japanische/Luxusyachten.</td>
</tr>
<tr>
<td>3</td>
<td>mm-sg</td>
<td>Den Ort/auf dem Hügel/eroberte/sieben Tage lang/eine gut bewaffnete/Armee.</td>
</tr>
<tr>
<td></td>
<td>cc-sg</td>
<td>Den Ort/auf dem Hügel/eroberte/sieben Tage/eine gut bewaffnete/Armee.</td>
</tr>
<tr>
<td></td>
<td>ctrl-sg</td>
<td>Den Ort/auf dem Hügel/eroberte/vor sieben Tagen/eine gut bewaffnete/Armee.</td>
</tr>
<tr>
<td></td>
<td>mm-pl</td>
<td>Den Ort/auf dem Hügel/eroberten/sieben Tage lang/viele gut bewaffnete/Soldaten.</td>
</tr>
<tr>
<td>item</td>
<td>condition</td>
<td>sentence</td>
</tr>
<tr>
<td>------</td>
<td>-----------</td>
<td>----------</td>
</tr>
<tr>
<td>3</td>
<td>cc-pl</td>
<td>Den Ort/auf dem Hügel/eroberten/in sieben Tagen/viele gut bewaffnete/Soldaten.</td>
</tr>
<tr>
<td></td>
<td>ctrl-pl</td>
<td>Den Ort/auf dem Hügel/eroberten/vor sieben Tagen/viele gut bewaffnete/Soldaten.</td>
</tr>
<tr>
<td>4</td>
<td>mm-sg</td>
<td>Den Regimegegner/mit Kontakten ins Ausland/exekutierte/eine Stunde lang/ein militärisches/Erschießungskommando.</td>
</tr>
<tr>
<td></td>
<td>mm-pl</td>
<td>Den Regimegegner/mit Kontakten ins Ausland/exekutierten/eine Stunde lang/mehrere Handlanger/der Militärregierung.</td>
</tr>
<tr>
<td>5</td>
<td>mm-sg</td>
<td>Den Grund/für den Ehekrach/erkannte/eine halbe Stunde lang/ein beratender/Psychologe.</td>
</tr>
<tr>
<td></td>
<td>cc-sg</td>
<td>Den Grund/für den Ehekrach/erkannte/in einer halben Stunde/ein beratender/Psychologe.</td>
</tr>
<tr>
<td></td>
<td>ctrl-sg</td>
<td>Den Grund/für den Ehekrach/erkannte/vor einer halben Stunde/ein beratender/Psychologe.</td>
</tr>
<tr>
<td></td>
<td>mm-pl</td>
<td>Den Grund/für den Ehekrach/erkannten/eine halbe Stunde lang/die beiden beratenden/Psychologen.</td>
</tr>
<tr>
<td></td>
<td>cc-pl</td>
<td>Den Grund/für den Ehekrach/erkannten/in einer halben Stunde/die beiden beratenden/Psychologen.</td>
</tr>
<tr>
<td></td>
<td>ctrl-pl</td>
<td>Den Grund/für den Ehekrach/erkannten/vor einer halben Stunde/die beiden beratenden/Psychologen.</td>
</tr>
<tr>
<td>6</td>
<td>mm-sg</td>
<td>Den Bankräuber/mit den vielen Vorstrafen/überführte/eine Stunde lang/ein äußerst gewitzter/Kommissar.</td>
</tr>
<tr>
<td></td>
<td>cc-sg</td>
<td>Den Bankräuber/mit den vielen Vorstrafen/überführte/in einer Stunde/ein äußerst gewitzter/Kommissar.</td>
</tr>
<tr>
<td></td>
<td>ctrl-sg</td>
<td>Den Bankräuber/mit den vielen Vorstrafen/überführte/vor einer Stunde/ein äußerst gewitzter/Kommissar.</td>
</tr>
<tr>
<td></td>
<td>mm-pl</td>
<td>Den Bankräuber/mit den vielen Vorstrafen/überführten/eine Stunde lang/zwei äußerst gewitzte/Kommissare.</td>
</tr>
<tr>
<td></td>
<td>cc-pl</td>
<td>Den Bankräuber/mit den vielen Vorstrafen/überführten/in einer Stunde/zwei äußerst gewitzte/Kommissare.</td>
</tr>
<tr>
<td></td>
<td>ctrl-pl</td>
<td>Den Bankräuber/mit den vielen Vorstrafen/überführten/vor einer Stunde/zwei äußerst gewitzte/Kommissare.</td>
</tr>
<tr>
<td>7</td>
<td>mm-sg</td>
<td>Den Unterschlupf/des berüchtigten Wilderers/entdeckte/eine halbe Stunde lang/ein aufmerksamer/läger.</td>
</tr>
<tr>
<td></td>
<td>cc-sg</td>
<td>Den Unterschlupf/des berüchtigten Wilderers/entdeckte/in einer halben Stunde/ein aufmerksamer/läger.</td>
</tr>
</tbody>
</table>

Appendix J. Target sentences in Experiment 8

<table>
<thead>
<tr>
<th>item</th>
<th>condition</th>
<th>sentence</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>ctrl-sg</td>
<td>Den Unterschlupf/des berüchtigten Wilderers/entdeckte/vor einer halben Stunde/ein aufmerksamer/läger.</td>
</tr>
<tr>
<td></td>
<td>mm-pl</td>
<td>Den Unterschlupf/des berüchtigten Wilderers/entdeckten/eine halbe Stunde lang/mehrere aufmerksame/läger.</td>
</tr>
<tr>
<td></td>
<td>cc-pl</td>
<td>Den Unterschlupf/des berüchtigten Wilderers/entdeckten/in einer halben Stunde/mehrere aufmerksame/läger.</td>
</tr>
<tr>
<td></td>
<td>ctrl-pl</td>
<td>Den Unterschlupf/des berüchtigten Wilderers/entdeckten/vor einer halben Stunde/mehrere aufmerksame/läger.</td>
</tr>
<tr>
<td>8</td>
<td>mm-sg</td>
<td>Den Fehler/am Vergaser/identifizierte/zwanzig Minuten lang/ein geschickter/Mechaniker.</td>
</tr>
<tr>
<td></td>
<td>cc-sg</td>
<td>Den Fehler/am Vergaser/identifizierte/in zwanzig Minuten/ein geschickter/Mechaniker.</td>
</tr>
<tr>
<td></td>
<td>ctrl-sg</td>
<td>Den Fehler/am Vergaser/identifizierte/vor zwanzig Minuten/ein geschickter/Mechaniker.</td>
</tr>
<tr>
<td></td>
<td>mm-pl</td>
<td>Den Fehler/am Vergaser/identifizierten/zwanzig Minuten lang/ein paar geschickte/Mechaniker.</td>
</tr>
<tr>
<td></td>
<td>cc-pl</td>
<td>Den Fehler/am Vergaser/identifizierten/in zwanzig Minuten/ein paar geschickte/Mechaniker.</td>
</tr>
<tr>
<td></td>
<td>ctrl-pl</td>
<td>Den Fehler/am Vergaser/identifizierten/vor zwanzig Minuten/ein paar geschickte/Mechaniker.</td>
</tr>
<tr>
<td>9</td>
<td>mm-sg</td>
<td>Den Löwen/mit der prächtigen Mähne/erlegte/drei Stunden lang/ein südafrikanischer/Großwildjäger.</td>
</tr>
<tr>
<td></td>
<td>cc-sg</td>
<td>Den Löwen/mit der prächtigen Mähne/erlegte/in drei Stunden/ein südafrikanischer/Großwildjäger.</td>
</tr>
<tr>
<td></td>
<td>ctrl-sg</td>
<td>Den Löwen/mit der prächtigen Mähne/erlegte/vor drei Stunden/ein südafrikanischer/Großwildjäger.</td>
</tr>
<tr>
<td></td>
<td>mm-pl</td>
<td>Den Löwen/mit der prächtigen Mähne/erlegten/drei Stunden lang/mehrere südafrikanische/Großwildjäger.</td>
</tr>
<tr>
<td></td>
<td>cc-pl</td>
<td>Den Löwen/mit der prächtigen Mähne/erlegten/in drei Stunden/mehrere südafrikanische/Großwildjäger.</td>
</tr>
<tr>
<td></td>
<td>ctrl-pl</td>
<td>Den Löwen/mit der prächtigen Mähne/erlegten/vor drei Stunden/mehrere südafrikanische/Großwildjäger.</td>
</tr>
<tr>
<td>10</td>
<td>mm-sg</td>
<td>Den Bullen/von Bauer Friedrich/tötete/eine halbe Stunde lang/ein ansässiger/Tierarzt.</td>
</tr>
<tr>
<td></td>
<td>cc-sg</td>
<td>Den Bullen/von Bauer Friedrich/tötete/in einer halben Stunde/ein ansässiger/Tierarzt.</td>
</tr>
<tr>
<td></td>
<td>ctrl-sg</td>
<td>Den Bullen/von Bauer Friedrich/tötete/vor einer halben Stunde/ein ansässiger/Tierarzt.</td>
</tr>
<tr>
<td></td>
<td>mm-pl</td>
<td>Den Bullen/von Bauer Friedrich/töteten/eine halbe Stunde lang/zwei ansässige/Tierärzte.</td>
</tr>
<tr>
<td></td>
<td>cc-pl</td>
<td>Den Bullen/von Bauer Friedrich/töteten/in einer halben Stunde/zwei ansässige/Tierärzte.</td>
</tr>
<tr>
<td></td>
<td>ctrl-pl</td>
<td>Den Bullen/von Bauer Friedrich/töteten/vor einer halben Stunde/zwei ansässige/Tierärzte.</td>
</tr>
<tr>
<td>item</td>
<td>condition</td>
<td>sentence</td>
</tr>
<tr>
<td>------</td>
<td>-----------</td>
<td>----------</td>
</tr>
<tr>
<td>11</td>
<td>mm-sg</td>
<td>Das Zeugnis für die beste Dissertation erhielt/dreißig Minuten lang/der junge/Doktorand.</td>
</tr>
<tr>
<td></td>
<td>cc-sg</td>
<td>Das Zeugnis für die beste Dissertation erhielt/in dreißig Minuten/der junge/Doktorand.</td>
</tr>
<tr>
<td></td>
<td>ctrl-sg</td>
<td>Das Zeugnis für die beste Dissertation erhielt/vor dreißig Minuten/der junge/Doktorand.</td>
</tr>
<tr>
<td></td>
<td>mm-pl</td>
<td>Das Zeugnis für die beste Dissertation erhielten/dreißig Minuten lang/zwei junge/Doktoranden.</td>
</tr>
<tr>
<td></td>
<td>cc-pl</td>
<td>Das Zeugnis für die beste Dissertation erhielten/in dreißig Minuten/zwei junge/Doktoranden.</td>
</tr>
<tr>
<td></td>
<td>ctrl-pl</td>
<td>Das Zeugnis für die beste Dissertation erhielten/vor dreißig Minuten/zwei junge/Doktoranden.</td>
</tr>
<tr>
<td>12</td>
<td>mm-sg</td>
<td>Den Schatz im Wald fand/fünf Minuten lang/der kleine/Peter.</td>
</tr>
<tr>
<td></td>
<td>cc-sg</td>
<td>Den Schatz im Wald fand/in fünf Minuten/der kleine/Peter.</td>
</tr>
<tr>
<td></td>
<td>ctrl-sg</td>
<td>Den Schatz im Wald fand/vor fünf Minuten/der kleine/Peter.</td>
</tr>
<tr>
<td></td>
<td>mm-pl</td>
<td>Den Schatz im Wald fanden/fünf Minuten lang/mehrere eifrig suchende/Kinder.</td>
</tr>
<tr>
<td></td>
<td>cc-pl</td>
<td>Den Schatz im Wald fanden/in fünf Minuten/mehrere eifrig suchende/Kinder.</td>
</tr>
<tr>
<td></td>
<td>ctrl-pl</td>
<td>Den Schatz im Wald fanden/vor fünf Minuten/mehrere eifrig suchende/Kinder.</td>
</tr>
<tr>
<td>13</td>
<td>mm-sg</td>
<td>Den Text über Produktionsästhetik verstand/einen Nachmittag lang/sogar/ein Medizinstudent.</td>
</tr>
<tr>
<td></td>
<td>cc-sg</td>
<td>Den Text über Produktionsästhetik verstand/in einem Nachmittag/sogar/ein Medizinstudent.</td>
</tr>
<tr>
<td></td>
<td>ctrl-sg</td>
<td>Den Text über Produktionsästhetik verstand/vor einem Nachmittag/sogar/ein Medizinstudent.</td>
</tr>
<tr>
<td></td>
<td>mm-pl</td>
<td>Den Text über Produktionsästhetik verstanden/einen Nachmittag lang/sogar/zwei Medizinstudenten.</td>
</tr>
<tr>
<td></td>
<td>cc-pl</td>
<td>Den Text über Produktionsästhetik verstanden/in einem Nachmittag/sogar/zwei Medizinstudenten.</td>
</tr>
<tr>
<td></td>
<td>ctrl-pl</td>
<td>Den Text über Produktionsästhetik verstanden/vor einem Nachmittag/sogar/zwei Medizinstudenten.</td>
</tr>
<tr>
<td>14</td>
<td>mm-sg</td>
<td>Den Schacht im alten Bergwerk sprengte/zwei Tage lang/ein erfahrener/Sprengmeister.</td>
</tr>
<tr>
<td></td>
<td>cc-sg</td>
<td>Den Schacht im alten Bergwerk sprengte/in zwei Tagen/ein erfahrener/Sprengmeister.</td>
</tr>
<tr>
<td></td>
<td>ctrl-sg</td>
<td>Den Schacht im alten Bergwerk sprengte/vor zwei Tagen/ein erfahrener/Sprengmeister.</td>
</tr>
<tr>
<td></td>
<td>mm-pl</td>
<td>Den Schacht im alten Bergwerk sprengten/zwei Tage lang/mehrere erfahrene/Sprengmeister.</td>
</tr>
<tr>
<td>15</td>
<td>mm-sg</td>
<td>Den ersten selbstreinigenden Teppich erfand/fünf Tage lang/ein ukrainischer/Erfinder.</td>
</tr>
<tr>
<td></td>
<td>cc-sg</td>
<td>Den ersten selbstreinigenden Teppich erfand/in fünf Tagen/ein ukrainischer/Erfinder.</td>
</tr>
<tr>
<td></td>
<td>ctrl-sg</td>
<td>Den ersten selbstreinigenden Teppich erfand/vor fünf Tagen/ein ukrainischer/Erfinder.</td>
</tr>
<tr>
<td></td>
<td>mm-pl</td>
<td>Den ersten selbstreinigenden Teppich erfanden/fünf Tage lang/ein paar ukrainische/Erfinder.</td>
</tr>
<tr>
<td></td>
<td>cc-pl</td>
<td>Den ersten selbstreinigenden Teppich erfanden/in fünf Tagen/ein paar ukrainische/Erfinder.</td>
</tr>
<tr>
<td></td>
<td>ctrl-pl</td>
<td>Den ersten selbstreinigenden Teppich erfanden/vor fünf Tagen/ein paar ukrainische/Erfinder.</td>
</tr>
<tr>
<td>16</td>
<td>mm-sg</td>
<td>Den Anruf über die Firmenpleite bekam/fünf Minuten lang/ein konkurrierender/Unternehmer.</td>
</tr>
<tr>
<td></td>
<td>cc-sg</td>
<td>Den Anruf über die Firmenpleite bekam/in fünf Minuten/ein konkurrierender/Unternehmer.</td>
</tr>
<tr>
<td></td>
<td>ctrl-sg</td>
<td>Den Anruf über die Firmenpleite bekam/vor fünf Minuten/ein konkurrierender/Unternehmer.</td>
</tr>
<tr>
<td></td>
<td>mm-pl</td>
<td>Den Anruf über die Firmenpleite bekamen/fünf Minuten lang/alle konkurrierenden/Unternehmer.</td>
</tr>
<tr>
<td></td>
<td>cc-pl</td>
<td>Den Anruf über die Firmenpleite bekamen/in fünf Minuten/alle konkurrierenden/Unternehmer.</td>
</tr>
<tr>
<td></td>
<td>ctrl-pl</td>
<td>Den Anruf über die Firmenpleite bekamen/vor fünf Minuten/alle konkurrierenden/Unternehmer.</td>
</tr>
<tr>
<td>17</td>
<td>mm-sg</td>
<td>Den Haarriss am Wasserrohr bemerkte/dreißig Minuten lang/ein aufmerksamer/Klempner.</td>
</tr>
<tr>
<td></td>
<td>cc-sg</td>
<td>Den Haarriss am Wasserrohr bemerkte/in dreißig Minuten/ein aufmerksamer/Klempner.</td>
</tr>
<tr>
<td></td>
<td>ctrl-sg</td>
<td>Den Haarriss am Wasserrohr bemerkte/vor dreißig Minuten/ein aufmerksamer/Klempner.</td>
</tr>
<tr>
<td></td>
<td>mm-pl</td>
<td>Den Haarriss am Wasserrohr bemerkten/dreißig Minuten lang/zwei aufmerksamere/Klempner.</td>
</tr>
<tr>
<td></td>
<td>cc-pl</td>
<td>Den Haarriss am Wasserrohr bemerkten/in dreißig Minuten/zwei aufmerksamere/Klempner.</td>
</tr>
<tr>
<td></td>
<td>ctrl-pl</td>
<td>Den Haarriss am Wasserrohr bemerkten/vor dreißig Minuten/zwei aufmerksamere/Klempner.</td>
</tr>
<tr>
<td>18</td>
<td>mm-sg</td>
<td>Den Höhleingang mit dem engen Durchschlupf verließ/fünfzehn Minuten lang/ein etwas dickerer/Höhlenforscher.</td>
</tr>
<tr>
<td></td>
<td>cc-sg</td>
<td>Den Höhleingang mit dem engen Durchschlupf verließ/in fünfzehn Minuten/ein etwas dickerer/Höhlenforscher.</td>
</tr>
<tr>
<td>item</td>
<td>condition</td>
<td>sentence</td>
</tr>
<tr>
<td>------</td>
<td>-----------</td>
<td>----------</td>
</tr>
<tr>
<td>18</td>
<td>ctrl-sg</td>
<td>Den Höhleneingang/mit dem engen Durchschlupf/verließ/vor fünfzehn Minuten/ein etwas dickere/Höhlforscher.</td>
</tr>
<tr>
<td></td>
<td>mm-pl</td>
<td>Den Höhleneingang/mit dem engen Durchschlupf/verließen/in fünfzehn Minuten/lang/ein paar etwas dickere/Höhlforscher.</td>
</tr>
<tr>
<td>19</td>
<td>mm-sg</td>
<td>Den Aufgabenteil/mit der Kurvendiskussion/bewältigte/eine Viertel Stunde/lang/ein äußerst begabter/Grundschüler.</td>
</tr>
<tr>
<td></td>
<td>ctrl-sg</td>
<td>Den Aufgabenteil/mit der Kurvendiskussion/bewältigte/vor einer Viertel Stunde/ein äußerst begabter/Grundschüler.</td>
</tr>
<tr>
<td></td>
<td>mm-pl</td>
<td>Den Aufgabenteil/mit der Kurvendiskussion/bewältigten/in einer Viertel Stunde/mehrere äußerst begabte/Grundschüler.</td>
</tr>
<tr>
<td>20</td>
<td>mm-sg</td>
<td>Den Beweis/des komplizierten Theorems/durchschüttete/zwei Tage/lang/ein daran arbeitender/Mathematik-Professor.</td>
</tr>
<tr>
<td></td>
<td>cc-sg</td>
<td>Den Beweis/des komplizierten Theorems/durchschüttete/in zwei Tagen/ein daran arbeitender/Mathematik-Professor.</td>
</tr>
<tr>
<td></td>
<td>ctrl-sg</td>
<td>Den Beweis/des komplizierten Theorems/durchschüttete/vor zwei Tagen/ein daran arbeitender/Mathematik-Professor.</td>
</tr>
<tr>
<td></td>
<td>mm-pl</td>
<td>Den Beweis/des komplizierten Theorems/durchschütteten/zwei Tage/lang/zwei daran arbeitende/Mathematik-Professorinnen.</td>
</tr>
<tr>
<td></td>
<td>cc-pl</td>
<td>Den Beweis/des komplizierten Theorems/durchschütteten/in zwei Tagen/zwei daran arbeitende/Mathematik-Professorinnen.</td>
</tr>
<tr>
<td></td>
<td>ctrl-pl</td>
<td>Den Beweis/des komplizierten Theorems/durchschütteten/vor zwei Tagen/zwei daran arbeitende/Mathematik-Professorinnen.</td>
</tr>
<tr>
<td>21</td>
<td>mm-sg</td>
<td>Den Raumgleiter/zum Mars/startete/eine Stunde/lang/die amerikanische/Raumfahrtbehörde.</td>
</tr>
<tr>
<td></td>
<td>cc-sg</td>
<td>Den Raumgleiter/zum Mars/startete/in einer Stunde/die amerikanische/Raumfahrtbehörde.</td>
</tr>
<tr>
<td></td>
<td>ctrl-sg</td>
<td>Den Raumgleiter/zum Mars/startete/vor einer Stunde/die amerikanische/Raumfahrtbehörde.</td>
</tr>
<tr>
<td></td>
<td>mm-pl</td>
<td>Den Raumgleiter/zum Mars/starteten/eine Stunde/lang/mehrere amerikanische/Astronauten.</td>
</tr>
<tr>
<td></td>
<td>cc-pl</td>
<td>Den Raumgleiter/zum Mars/starteten/in einer Stunde/mehrere amerikanische/Astronauten.</td>
</tr>
<tr>
<td></td>
<td>ctrl-pl</td>
<td>Den Raumgleiter/zum Mars/starteten/vor einer Stunde/mehrere amerikanische/Astronauten.</td>
</tr>
<tr>
<td>22</td>
<td>mm-sg</td>
<td>Den Patienten/aus Sachsen/befielen/vierundzwanzig Stunden/lang/eine gefährliche/Tropenkrankheit.</td>
</tr>
<tr>
<td></td>
<td>cc-sg</td>
<td>Den Patienten/aus Sachsen/befielen/in vierundzwanzig Stunden/eine gefährliche/Tropenkrankheit.</td>
</tr>
<tr>
<td></td>
<td>ctrl-sg</td>
<td>Den Patienten/aus Sachsen/befielen/vor vierundzwanzig Stunden/eine gefährliche/Tropenkrankheit.</td>
</tr>
<tr>
<td></td>
<td>mm-pl</td>
<td>Den Patienten/aus Sachsen/befielen/vierundzwanzig Stunden/lang/äußert gefährliche/Viren.</td>
</tr>
<tr>
<td></td>
<td>cc-pl</td>
<td>Den Patienten/aus Sachsen/befielen/in vierundzwanzig Stunden/äußert gefährliche/Viren.</td>
</tr>
<tr>
<td></td>
<td>ctrl-pl</td>
<td>Den Patienten/aus Sachsen/befielen/vor vierundzwanzig Stunden/äußert gefährliche/Viren.</td>
</tr>
<tr>
<td>23</td>
<td>mm-sg</td>
<td>Den Gegner/aus Russland/übermannte/eine halbe Stunde/lang/ein georgischer/Nachwuchsringen.</td>
</tr>
<tr>
<td></td>
<td>ctrl-sg</td>
<td>Den Gegner/aus Russland/übermannte/vor einer halben Stunde/ein georgischer/Nachwuchsringen.</td>
</tr>
<tr>
<td></td>
<td>mm-pl</td>
<td>Den Gegner/aus Russland/übermannten/in einer halben Stunde/lang/mehrere georgische/Nachwuchsringen.</td>
</tr>
<tr>
<td>24</td>
<td>mm-sg</td>
<td>Den Konkurrenten/vom McLaren-Team/besiegten/dreißig Minuten/lang/ein bislang unbekannter/Rennfahrer.</td>
</tr>
<tr>
<td></td>
<td>cc-sg</td>
<td>Den Konkurrenten/vom McLaren-Team/besiegte/in dreißig Minuten/ein bislang unbekannter/Rennfahrer.</td>
</tr>
<tr>
<td></td>
<td>ctrl-sg</td>
<td>Den Konkurrenten/vom McLaren-Team/besiegte/vor dreißig Minuten/ein bislang unbekannter/Rennfahrer.</td>
</tr>
<tr>
<td></td>
<td>mm-pl</td>
<td>Den Konkurrenten/vom McLaren-Team/besiegten/dreißig Minuten/lang/zwei eher unbekannte/Rennfahrer.</td>
</tr>
<tr>
<td></td>
<td>cc-pl</td>
<td>Den Konkurrenten/vom McLaren-Team/besiegten/in dreißig Minuten/zwei eher unbekannte/Rennfahrer.</td>
</tr>
<tr>
<td></td>
<td>ctrl-pl</td>
<td>Den Konkurrenten/vom McLaren-Team/besiegten/vor dreißig Minuten/zwei eher unbekannte/Rennfahrer.</td>
</tr>
<tr>
<td>25</td>
<td>mm-sg</td>
<td>Den Leuchtturm/auf der kleinen Insel/erspäht/einfünf Minuten/lang/alle Wache schiebende/Matrosen.</td>
</tr>
<tr>
<td></td>
<td>cc-sg</td>
<td>Den Leuchtturm/auf der kleinen Insel/erspähte/in fünf Minuten/ein Wache schiebender/Matrosen.</td>
</tr>
<tr>
<td></td>
<td>ctrl-sg</td>
<td>Den Leuchtturm/auf der kleinen Insel/erspähte/vor fünf Minuten/ein Wache schiebender/Matrosen.</td>
</tr>
<tr>
<td></td>
<td>mm-pl</td>
<td>Den Leuchtturm/auf der kleinen Insel/erspähen/fünf Minuten/lang/ein Wache schiebender/Matrosen.</td>
</tr>
<tr>
<td>item</td>
<td>condition</td>
<td>sentence</td>
</tr>
<tr>
<td>------</td>
<td>-----------</td>
<td>--------------------------------------------------------------------------</td>
</tr>
<tr>
<td>25</td>
<td>cc-pl</td>
<td>Den Leuchtturm/auf der kleinen Insel/erspähten/in fünf Minuten/alle Wache schiebenden/Matrosen.</td>
</tr>
<tr>
<td></td>
<td>ctrl-pl</td>
<td>Den Leuchtturm/auf der kleinen Insel/erspähten/vor fünf Minuten/alle Wache schiebenden/Matrosen.</td>
</tr>
<tr>
<td></td>
<td>mm-sg</td>
<td>Den Kollegen/aus Luxemburg/erblickte/zehn Minuten lang/ein deutscher/Konferenzzteilnehmer.</td>
</tr>
<tr>
<td></td>
<td>cc-sg</td>
<td>Den Kollegen/aus Luxemburg/erblickte/in zehn Minuten/ein deutscher/Konferenzzteilnehmer.</td>
</tr>
<tr>
<td></td>
<td>ctrl-sg</td>
<td>Den Kollegen/aus Luxemburg/erblickte/vor zehn Minuten/ein deutscher/Konferenzzteilnehmer.</td>
</tr>
<tr>
<td></td>
<td>mm-pl</td>
<td>Den Kollegen/aus Luxemburg/erblickten/zehn Minuten lang/ein paar deutsche/Konferenzzteilnehmer.</td>
</tr>
<tr>
<td></td>
<td>cc-pl</td>
<td>Den Kollegen/aus Luxemburg/erblickten/in zehn Minuten/ein paar deutsche/Konferenzzteilnehmer.</td>
</tr>
<tr>
<td></td>
<td>ctrl-pl</td>
<td>Den Kollegen/aus Luxemburg/erblickten/vor zehn Minuten/ein paar deutsche/Konferenzzteilnehmer.</td>
</tr>
<tr>
<td>26</td>
<td>mm-sg</td>
<td>Den Lehrer/für Mathematik/verblüffte/zwanzig Minuten lang/eine äußerst begabte/Schülerin.</td>
</tr>
<tr>
<td></td>
<td>cc-sg</td>
<td>Den Lehrer/für Mathematik/verblüffte/in zwanzig Minuten/eine äußerst begabte/Schülerin.</td>
</tr>
<tr>
<td></td>
<td>ctrl-sg</td>
<td>Den Lehrer/für Mathematik/verblüffte/vor zwanzig Minuten/eine äußerst begabte/Schülerin.</td>
</tr>
<tr>
<td></td>
<td>mm-pl</td>
<td>Den Lehrer/für Mathematik/verblüffen/zwanzig Minuten lang/zwei äußerst begabte/Schülerinnen.</td>
</tr>
<tr>
<td></td>
<td>cc-pl</td>
<td>Den Lehrer/für Mathematik/verblüffen/in zwanzig Minuten/zwei äußerst begabte/Schülerinnen.</td>
</tr>
<tr>
<td></td>
<td>ctrl-pl</td>
<td>Den Lehrer/für Mathematik/verblüffen/vor zwanzig Minuten/zwei äußerst begabte/Schülerinnen.</td>
</tr>
<tr>
<td>27</td>
<td>mm-sg</td>
<td>Den Jungen/aus dem Nachbarhaus/erwischte/drei Minuten lang/der wütende/Hausmeister.</td>
</tr>
<tr>
<td></td>
<td>cc-sg</td>
<td>Den Jungen/aus dem Nachbarhaus/erwischte/in drei Minuten/der wütende/Hausmeister.</td>
</tr>
<tr>
<td></td>
<td>ctrl-sg</td>
<td>Den Jungen/aus dem Nachbarhaus/erwischte/vor drei Minuten/der wütende/Hausmeister.</td>
</tr>
<tr>
<td></td>
<td>mm-pl</td>
<td>Den Jungen/aus dem Nachbarhaus/erwischten/drei Minuten lang/die wütenden/Nachbarn.</td>
</tr>
<tr>
<td></td>
<td>cc-pl</td>
<td>Den Jungen/aus dem Nachbarhaus/erwischten/in drei Minuten/die wütenden/Nachbarn.</td>
</tr>
<tr>
<td></td>
<td>ctrl-pl</td>
<td>Den Jungen/aus dem Nachbarhaus/erwischten/vor drei Minuten/die wütenden/Nachbarn.</td>
</tr>
<tr>
<td>28</td>
<td>mm-sg</td>
<td>Den Anfang/des Arguments/vergaß/fünf Minuten lang/ein etwas zerstreuter/Redner.</td>
</tr>
<tr>
<td></td>
<td>cc-sg</td>
<td>Den Anfang/des Arguments/vergaß/in fünf Minuten/ein etwas zerstreuter/Redner.</td>
</tr>
<tr>
<td>29</td>
<td>ctrl-sg</td>
<td>Den Anfang/des Arguments/vergaß/vor fünf Minuten/die beiden zerstreuten/Redner.</td>
</tr>
<tr>
<td></td>
<td>mm-pl</td>
<td>Den Anfang/des Arguments/vergaß/fünf Minuten lang/die beiden zerstreuten/Redner.</td>
</tr>
<tr>
<td></td>
<td>cc-pl</td>
<td>Den Anfang/des Arguments/vergaß/in fünf Minuten/die beiden zerstreuten/Redner.</td>
</tr>
<tr>
<td></td>
<td>ctrl-pl</td>
<td>Den Anfang/des Arguments/vergaß/vor fünf Minuten/die beiden zerstreuten/Redner.</td>
</tr>
<tr>
<td>30</td>
<td>mm-sg</td>
<td>Den Ableger/der Stammfirma/gründete/zwei Jahre lang/ein geschickt kalkulierender/Manager.</td>
</tr>
<tr>
<td></td>
<td>cc-sg</td>
<td>Den Ableger/der Stammfirma/gründete/in zwei Jahren/ein geschickt kalkulierender/Manager.</td>
</tr>
<tr>
<td></td>
<td>ctrl-sg</td>
<td>Den Ableger/der Stammfirma/gründete/vor zwei Jahren/ein geschickt kalkulierender/Manager.</td>
</tr>
<tr>
<td></td>
<td>mm-pl</td>
<td>Den Ableger/der Stammfirma/gründeten/zwei Jahre lang/ein paar geschickt kalkulierende/Manager.</td>
</tr>
<tr>
<td></td>
<td>cc-pl</td>
<td>Den Ableger/der Stammfirma/gründeten/in zwei Jahren/ein paar geschickt kalkulierende/Manager.</td>
</tr>
<tr>
<td></td>
<td>ctrl-pl</td>
<td>Den Ableger/der Stammfirma/gründeten/vor zwei Jahren/ein paar geschickt kalkulierende/Manager.</td>
</tr>
</tbody>
</table>

Note: mm = mismatch; cc = coercion; ctrl = control; sg = singular; pl = plural
APPENDIX K

Target sentences in Experiment 9

Table K.1. “Aktionsart”-Items

<table>
<thead>
<tr>
<th>item</th>
<th>condition</th>
<th>sentence</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>coercion</td>
<td>In/drei/Stunden/hatte/der/Fahrer/das/Rennen/gewonnen/obwohl/die Strecke/zum Teil/gefroren/war.</td>
</tr>
<tr>
<td></td>
<td>control</td>
<td>Vor/drei/Stunden/hatte/der/Fahrer/das/Rennen/gewonnen/obwohl/die Strecke/zum Teil/gefroren/war.</td>
</tr>
<tr>
<td>3</td>
<td>mismatch</td>
<td>Eine/halbe/Stunde/hatte/der/Teilnehmer/das/Spiel/gewonnen/obwohl/die Gegner/gegen ihn/zusammen/hielten.</td>
</tr>
<tr>
<td></td>
<td>coercion</td>
<td>In einer/halben/Stunde/hatte/der/Teilnehmer/das/Spiel/gewonnen/obwohl/die Gegner/gegen ihn/zusammen/hielten.</td>
</tr>
<tr>
<td></td>
<td>control</td>
<td>Vor einer/halben/Stunde/hatte/der/Teilnehmer/das/Spiel/gewonnen/obwohl/die Gegner/gegen ihn/zusammen/hielten.</td>
</tr>
<tr>
<td>4</td>
<td>mismatch</td>
<td>Eine/halbe/Stunde/hatte/die/Yacht/den/Hafen/erreicht/obwohl/kein/starker/Wind/blies.</td>
</tr>
<tr>
<td></td>
<td>coercion</td>
<td>In einer/halben/Stunde/hatte/die/Yacht/den/Hafen/erreicht/obwohl/kein/starker/Wind/blies.</td>
</tr>
<tr>
<td></td>
<td>control</td>
<td>Vor einer/halben/Stunde/hatte/die/Yacht/den/Hafen/erreicht/obwohl/kein/starker/Wind/blies.</td>
</tr>
<tr>
<td></td>
<td>coercion</td>
<td>In einer/halben/Stunde/hatte/der/Zug/den/Bahnhof/erreicht/obwohl/unterwegs/der Strom/ausgefallen/war.</td>
</tr>
<tr>
<td></td>
<td>control</td>
<td>Vor einer/halben/Stunde/hatte/der/Zug/den/Bahnhof/erreicht/obwohl/unterwegs/der Strom/ausgefallen/war.</td>
</tr>
<tr>
<td>item</td>
<td>condition</td>
<td>sentence</td>
</tr>
<tr>
<td>------</td>
<td>-----------</td>
<td>----------</td>
</tr>
<tr>
<td>6</td>
<td>mismatch</td>
<td>Vor einer halben Stunde hatte der Bote die Stadt erreicht, obwohl er den Weg nicht/gut kannte.</td>
</tr>
<tr>
<td></td>
<td>coercion</td>
<td>In einer halben Stunde hatte der Bote die Stadt erreicht, obwohl er den Weg nicht/gut kannte.</td>
</tr>
<tr>
<td></td>
<td>control</td>
<td>Vor einer halben Stunde hatte der Bote die Stadt erreicht, obwohl er den Weg nicht/gut kannte.</td>
</tr>
<tr>
<td>7</td>
<td>mismatch</td>
<td>Ganze zwei Tage hatten die Reiter das Tal erobert, obwohl die Mauern schwer zu überwinden waren.</td>
</tr>
<tr>
<td></td>
<td>coercion</td>
<td>In zwei Tagen hatten die Reiter das Tal erobert, obwohl die Mauern schwer zu überwinden waren.</td>
</tr>
<tr>
<td></td>
<td>control</td>
<td>Vor zwei Tagen hatten die Reiter das Tal erobert, obwohl die Mauern schwer zu überwinden waren.</td>
</tr>
<tr>
<td>8</td>
<td>mismatch</td>
<td>Ganze drei Stunden hatte die Jury die Sprecherin ernannt, obwohl die Mitglieder verschiedener Meinung waren.</td>
</tr>
<tr>
<td></td>
<td>coercion</td>
<td>In drei Stunden hatte die Jury die Sprecherin ernannt, obwohl die Mitglieder verschiedener Meinung waren.</td>
</tr>
<tr>
<td></td>
<td>control</td>
<td>Vor drei Stunden hatte die Jury die Sprecherin ernannt, obwohl die Mitglieder verschiedener Meinung waren.</td>
</tr>
<tr>
<td>9</td>
<td>mismatch</td>
<td>Ganze zwei Stunden hatte der Unternehmer den Nachfolger ernannt, obwohl mehrere Kandidaten zur Auswahl standen.</td>
</tr>
<tr>
<td></td>
<td>coercion</td>
<td>In zwei Stunden hatte der Unternehmer den Nachfolger ernannt, obwohl mehrere Kandidaten zur Auswahl standen.</td>
</tr>
<tr>
<td></td>
<td>control</td>
<td>Vor zwei Stunden hatte der Unternehmer den Nachfolger ernannt, obwohl mehrere Kandidaten zur Auswahl standen.</td>
</tr>
<tr>
<td>10</td>
<td>mismatch</td>
<td>Ganze zehn Minuten hatte die Klasse den Klassensprecher ernannt, obwohl sich kein Schüler freiwillig gemeldet hatte.</td>
</tr>
<tr>
<td></td>
<td>coercion</td>
<td>In zehn Minuten hatte die Klasse den Klassensprecher ernannt, obwohl sich kein Schüler freiwillig gemeldet hatte.</td>
</tr>
<tr>
<td></td>
<td>control</td>
<td>Vor zehn Minuten hatte die Klasse den Klassensprecher ernannt, obwohl sich kein Schüler freiwillig gemeldet hatte.</td>
</tr>
</tbody>
</table>
### Appendix K. Target sentences in Experiment 9

<table>
<thead>
<tr>
<th>item</th>
<th>condition</th>
<th>sentence</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>mismatch</td>
<td>Ganze/zehn/Minuten/hatte/das/Kind/die/Ostereier/entdeckt/obwohl/der/Garten/riesig/war.</td>
</tr>
<tr>
<td></td>
<td>coercion</td>
<td>In/zehn/Minuten/hatte/das/Kind/die/Ostereier/entdeckt/obwohl/der/Garten/riesig/war.</td>
</tr>
<tr>
<td></td>
<td>control</td>
<td>Vor/zehn/Minuten/hatte/das/Kind/die/Ostereier/entdeckt/obwohl/der/Garten/riesig/war.</td>
</tr>
<tr>
<td></td>
<td>coercion</td>
<td>In/zwei/Wochen/hatte/der/Archäologe/die/Skulptur/entdeckt/obwohl/das/Grabmal/teilweise/verschüttet/war.</td>
</tr>
<tr>
<td></td>
<td>control</td>
<td>Vor/zwei/Wochen/hatte/der/Archäologe/die/Skulptur/entdeckt/obwohl/das/Grabmal/teilweise/verschüttet/war.</td>
</tr>
<tr>
<td></td>
<td>coercion</td>
<td>In/zwanzig/Minuten/hatte/der/Mechaniker/den/Schaden/identifiziert/obwohl/der Motor/nahezu/unzugänglich/war.</td>
</tr>
<tr>
<td></td>
<td>control</td>
<td>Vor/zwanzig/Minuten/hatte/der/Mechaniker/den/Schaden/identifiziert/obwohl/der Motor/nahezu/unzugänglich/war.</td>
</tr>
<tr>
<td>23</td>
<td>mismatch</td>
<td>Ganze/zwanzig/Minuten/hatte/der/Biologe/die/Pflanze/identifiziert/obwohl/sie/von Tieren/angefressen/war.</td>
</tr>
<tr>
<td></td>
<td>coercion</td>
<td>In/zwanzig/Minuten/hatte/der/Biologe/die/Pflanze/identifiziert/obwohl/sie/von Tieren/angefressen/war.</td>
</tr>
<tr>
<td></td>
<td>control</td>
<td>Vor/zwanzig/Minuten/hatte/der/Biologe/die/Pflanze/identifiziert/obwohl/sie/von Tieren/angefressen/war.</td>
</tr>
<tr>
<td>24</td>
<td>mismatch</td>
<td>Ganze/zwei/Stunden/hatte/der/Pathologe/die/Leiche/identifiziert/obwohl/sie/überall/Verbrünungen/hatte.</td>
</tr>
<tr>
<td></td>
<td>coercion</td>
<td>In/zwei/Stunden/hatte/der/Pathologe/die/Leiche/identifiziert/obwohl/sie/überall/Verbrünungen/hatte.</td>
</tr>
<tr>
<td></td>
<td>control</td>
<td>Vor/zwei/Stunden/hatte/der/Pathologe/die/Leiche/identifiziert/obwohl/sie/überall/Verbrünungen/hatte.</td>
</tr>
<tr>
<td></td>
<td>coercion</td>
<td>In/zwanzig/Minuten/hatte/der/Jäger/das/Reh/erlegt/obwohl es/nicht/auf/die Lichtung/kam.</td>
</tr>
<tr>
<td></td>
<td>control</td>
<td>Vor/zwanzig/Minuten/hatte/der/Jäger/das/Reh/erlegt/obwohl es/nicht/auf/die Lichtung/kam.</td>
</tr>
<tr>
<td></td>
<td>coercion</td>
<td>In/zwanzig/Minuten/hatte/der/Wilderer/den/Hirsch/erlegt/obwohl/drei/Kugeln/daneben/gingen.</td>
</tr>
<tr>
<td></td>
<td>control</td>
<td>Vor/zwanzig/Minuten/hatte/der/Wilderer/den/Hirsch/erlegt/obwohl/drei/Kugeln/daneben/gingen.</td>
</tr>
</tbody>
</table>
### Appendix K. Target sentences in Experiment 9

<table>
<thead>
<tr>
<th>item</th>
<th>condition</th>
<th>sentence</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>coercion</td>
<td>In/zwei/Tage/hatte/der/Junge/die/Geschenke/gefunden/obwohl/seine Mutter/sie/gut versteckt/hatte.</td>
</tr>
<tr>
<td></td>
<td>control</td>
<td>Vor/zwei/Tage/hatte/der/Junge/die/Geschenke/gefunden/obwohl/seine Mutter/sie/gut versteckt/hatte.</td>
</tr>
<tr>
<td></td>
<td>coercion</td>
<td>In/fünf/Minuten/hatte/der/Programmierer/den/Fehler/gefunden/obwohl/seine Kollegen/alle/aufgegeben/hatten.</td>
</tr>
<tr>
<td></td>
<td>control</td>
<td>Vor/fünf/Minuten/hatte/der/Programmierer/den/Fehler/gefunden/obwohl/seine Kollegen/alle/aufgegeben/hatten.</td>
</tr>
<tr>
<td>36</td>
<td>mismatch</td>
<td>Ganze/zwei/Stunden/hatte/die/Ehefrau/die/Liebesbriefe/gefunden/obwohl/ihr Mann/sie/gut versteckt/hatte.</td>
</tr>
<tr>
<td></td>
<td>coercion</td>
<td>In/zwei/Stunden/hatte/die/Ehefrau/die/Liebesbriefe/gefunden/obwohl/ihr Mann/sie/gut versteckt/hatte.</td>
</tr>
<tr>
<td></td>
<td>control</td>
<td>Vor/zwei/Stunden/hatte/die/Ehefrau/die/Liebesbriefe/gefunden/obwohl/ihr Mann/sie/gut versteckt/hatte.</td>
</tr>
<tr>
<td></td>
<td>coercion</td>
<td>In/zwei/Tagen/hatte/die/Studentin/den/Text/verstanden/obwohl/er/sehr/lang/war.</td>
</tr>
<tr>
<td></td>
<td>control</td>
<td>Vor/zwei/Tagen/hatte/die/Studentin/den/Text/verstanden/obwohl/er/sehr/lang/war.</td>
</tr>
<tr>
<td></td>
<td>coercion</td>
<td>In/zwei/Tagen/hatte/der/Übersetzer/den/Artikel/verstanden/obwohl/er/sehr/schwierig/war.</td>
</tr>
<tr>
<td></td>
<td>control</td>
<td>Vor/zwei/Tagen/hatte/der/Übersetzer/den/Artikel/verstanden/obwohl/er/sehr/schwierig/war.</td>
</tr>
</tbody>
</table>
In dreißig Minuten hatte der Regisseur den Einfall bekommen.

Ganze dreißig Minuten hatte der Regisseur den Einfall.

Vor fünfzehn Minuten hatte der Abenteurer die Heimatstadt verlassen.

Vor zehn Minuten hatte der Klempner den Rohrbruch bemerkt.

In fünfzehn Minuten hatte der Forscher die Höhle verlassen.

Ganze fünfzehn Minuten hatte die Botschafter das Land verlassen.

Vor fünfzehn Minuten hatte der Abenteurer die Heimatstadt aufgegeben.

In zwei Tagen hatte die Abenteurer den Ausflug des Großbreiten.

Vor zwei Tagen hatte der Forscher die Heimatstadt verlassen.

Ganze zwanzig Minuten hatte der Spezialist das Problem durchschaut.

Vor zwei Tagen hatte der Spezialist das Problem durchschaut.

Eine ganze Viertel Stunde hatte der Schüler die Aufgabe bewältigt.

In einer Viertel Stunde hatte der Schüler die Aufgabe bewältigt.
### Appendix K. Target sentences in Experiment 9

<table>
<thead>
<tr>
<th>item</th>
<th>condition</th>
<th>sentence</th>
</tr>
</thead>
<tbody>
<tr>
<td>62</td>
<td>mismatch</td>
<td>Eine/ganz/Stunde/hatte/der/Physiker/den/Versuch/gestartet/trotz/kleinerer/Pannen/beim/Versuchsaufbau.</td>
</tr>
<tr>
<td></td>
<td>coercion</td>
<td>In/einer/Stunde/hatte/der/Physiker/den/Versuch/gestartet/trotz/kleinerer/Pannen/beim/Versuchsaufbau.</td>
</tr>
<tr>
<td></td>
<td>control</td>
<td>Vor/einer/Stunde/hatte/der/Physiker/den/Versuch/gestartet/trotz/kleinerer/Pannen/beim/Versuchsaufbau.</td>
</tr>
<tr>
<td>63</td>
<td>mismatch</td>
<td>Eine/ganz/Stunde/hatte/das/Militär/die/Operation/gestartet/trotz/kleinerer/Unstimmigkeiten/im/Führungsstab.</td>
</tr>
<tr>
<td></td>
<td>coercion</td>
<td>In/einer/Stunde/hatte/das/Militär/die/Operation/gestartet/trotz/kleinerer/Unstimmigkeiten/im/Führungsstab.</td>
</tr>
<tr>
<td></td>
<td>control</td>
<td>Vor/einer/Stunde/hatte/das/Militär/die/Operation/gestartet/trotz/kleinerer/Unstimmigkeiten/im/Führungsstab.</td>
</tr>
<tr>
<td>64</td>
<td>mismatch</td>
<td>Ganze/vierundzwanzig/Stunden/hatte/das/Virus/den/Patienten/befallen/trotz/sorgfältiger/Maßnahmen/zur/Quarantäne.</td>
</tr>
<tr>
<td></td>
<td>coercion</td>
<td>In/vierundzwanzig/Stunden/hatte/das/Virus/den/Patienten/befallen/trotz/sorgfältiger/Maßnahmen/zur/Quarantäne.</td>
</tr>
<tr>
<td></td>
<td>control</td>
<td>Vor/vierundzwanzig/Stunden/hatte/das/Virus/den/Patienten/befallen/trotz/sorgfältiger/Maßnahmen/zur/Quarantäne.</td>
</tr>
<tr>
<td>65</td>
<td>mismatch</td>
<td>Ganze/vierundzwanzig/Stunden/hatten/die/Blattläuse/die/Pflanze/befallen/trotz/der/Verwendung/von/Pestiziden.</td>
</tr>
<tr>
<td></td>
<td>coercion</td>
<td>In/vierundzwanzig/Stunden/hatten/die/Blattläuse/die/Pflanze/befallen/trotz/der/Verwendung/von/Pestiziden.</td>
</tr>
<tr>
<td></td>
<td>control</td>
<td>Vor/vierundzwanzig/Stunden/hatten/die/Blattläuse/die/Pflanze/befallen/trotz/der/Verwendung/von/Pestiziden.</td>
</tr>
<tr>
<td></td>
<td>coercion</td>
<td>In/vierundzwanzig/Stunden/hatte/die/Faulnis/die/Ernte/befallen/obwohl/der/Bauer/diese/zu/trocknen/versuchte.</td>
</tr>
<tr>
<td></td>
<td>control</td>
<td>Vor/vierundzwanzig/Stunden/hatte/die/Faulnis/die/Ernte/befallen/obwohl/der/Bauer/diese/zu/trocknen/versuchte.</td>
</tr>
<tr>
<td></td>
<td>coercion</td>
<td>In/einer/halben/Stunde/hatte/der/Boxer/den/Gegner/übermannt/trotz/mehrerer/knapp/verlorenen/Runden.</td>
</tr>
<tr>
<td></td>
<td>control</td>
<td>Vor/einer/halben/Stunde/hatte/der/Boxer/den/Gegner/übermannt/trotz/mehrerer/knapp/verlorenen/Runden.</td>
</tr>
<tr>
<td>68</td>
<td>mismatch</td>
<td>Eine/halbe/Stunde/hatte/die/Angst/den/Trapezkünstler/übermannt/obwohl/er/nie/gefährlich/stürzte.</td>
</tr>
<tr>
<td></td>
<td>coercion</td>
<td>In/einer/halben/Stunde/hatte/die/Angst/den/Trapezkünstler/übermannt/obwohl/er/nie/gefährlich/stürzte.</td>
</tr>
<tr>
<td></td>
<td>control</td>
<td>Vor/einer/halben/Stunde/hatte/die/Angst/den/Trapezkünstler/übermannt/obwohl/er/nie/gefährlich/stürzte.</td>
</tr>
</tbody>
</table>
### Appendix K. Target sentences in Experiment 9

<table>
<thead>
<tr>
<th>item</th>
<th>condition</th>
<th>sentence</th>
</tr>
</thead>
<tbody>
<tr>
<td>76</td>
<td>mismatch</td>
<td>Ganze/zehn/Minuten/hatte/der/Arzt/den/Kollegen/erblickt/obwohl/die/Tafel/etwas/versteckt/angebracht/ war.</td>
</tr>
<tr>
<td></td>
<td>coercion</td>
<td>In/zehn/Minuten/hatte/der/Arzt/den/Kollegen/erblickt/obwohl/die/Tafel/etwas/versteckt/angebracht/ war.</td>
</tr>
<tr>
<td></td>
<td>control</td>
<td>Vor/zehn/Minuten/hatte/der/Arzt/den/Kollegen/erblickt/obwohl/die/Tafel/etwas/versteckt/angebracht/ war.</td>
</tr>
<tr>
<td>77</td>
<td>mismatch</td>
<td>Ganze/zehn/Minuten/hatte/der/Hausfrau/das/Sonderangebot/erblickt/obwohl/die/Tafel/etwas/versteckt/angebracht/ war.</td>
</tr>
<tr>
<td></td>
<td>coercion</td>
<td>In/zehn/Minuten/hatte/der/Hausfrau/das/Sonderangebot/erblickt/obwohl/die/Tafel/etwas/versteckt/angebracht/ war.</td>
</tr>
<tr>
<td></td>
<td>control</td>
<td>Vor/zehn/Minuten/hatte/der/Hausfrau/das/Sonderangebot/erblickt/obwohl/die/Tafel/etwas/versteckt/angebracht/ war.</td>
</tr>
<tr>
<td></td>
<td>coercion</td>
<td>In/drei/Minuten/hatte/der/Schütze/das/Ziel/erfasst/obwohl/das/Zielfernrohr/nicht/richtig/funktionierte.</td>
</tr>
<tr>
<td></td>
<td>control</td>
<td>Vor/drei/Minuten/hatte/der/Schütze/das/Ziel/erfasst/obwohl/das/Zielfernrohr/nicht/richtig/funktionierte.</td>
</tr>
<tr>
<td></td>
<td>coercion</td>
<td>In/vierzig/Minuten/hatte/der/Studentin/den/Inhalt/erfasst/obwohl/der/Text/auf/Französisch/formuliert/ war.</td>
</tr>
<tr>
<td></td>
<td>control</td>
<td>Vor/vierzig/Minuten/hatte/der/Studentin/den/Inhalt/erfasst/obwohl/der/Text/auf/Französisch/formuliert/ war.</td>
</tr>
<tr>
<td></td>
<td>coercion</td>
<td>In/drei/Minuten/hatte/der/Hausmeister/den/Jungen/erwischt/obwohl/dieser/ihn/auszutricksen/versuchte.</td>
</tr>
<tr>
<td></td>
<td>control</td>
<td>Vor/drei/Minuten/hatte/der/Hausmeister/den/Jungen/erwischt/obwohl/dieser/ihn/auszutricksen/versuchte.</td>
</tr>
</tbody>
</table>
In fünf Minuten hatte der Sohn den Schlüssel vergessen / obwohl / eine Notiz auf dem Tisch / daran erinnerte.  
coercion: 
95  mismatch  Ganze fünf Minuten hatte / die / Großmutter / den / Heimweg / vergessen / trotz / der / detaillierten / Beschreibung / der / Tochter.  
coercion: 
coercion: 
91  mismatch  Eine ganze Minute hatte / die / Putzfrau / das / Fenster / geschlossen / obwohl / es / heftig / geklemmt / hatte.  
coercion: 
92  mismatch  Eine ganze Minute hatte / der / Busfahrer / die / Tür / geschlossen / obwohl / die Hydraulik / nicht / mehr / funktionierte.  
coercion: 
coercion: 
94  mismatch  Ganze fünf Minuten hatte / der / Mann / die / Telefonnummer / vergessen / obwohl / er / sie / zu / memorieren / versuchte.  
coercion: 
90  mismatch  Ganze zwei Wochen hatten / die / Studenten / die / Kampagne / gegründet / obwohl / die Unleitung / gegen / sie / agitierte.  
coercion: 
coercion: 
coercion: 
coercion: 
100  mismatch  Ganze zwei Monate hatte / der / Herausgeber / den / Text / akzeptiert / obwohl / die Verhandlung / mit / dem / Autor / schwierig war.  
coercion: 
101  mismatch  Ganze zwei Stunden hatte / der / Professor / den / Beweis / akzeptiert / obwohl er / anfangs / eine Prämisse / nicht / verstanden.  
coercion: 
102  mismatch  Ganze drei Tage hatte / die / Bank / den / Scheck / akzeptiert / trotz / des / anfänglichen / Verdachts / auf / Fälschung.  
coercion: 
103  mismatch  Ganze drei Tage hatte / die / Frau / die / Nachricht / erfahren / obwohl si / bereits / mehrfach zuvor / angerufen / hatte.  
coercion:
correction In/drei/Stunden/hatte/die/Sekretärin/die/Neuigkeit/erfahren/obwohl/sie/ständig/die Kollegen/ausbortchte.
control Vor/drei/Stunden/hatte/die/Sekretärin/die/Neuigkeit/erfahren/obwohl/sie/ständig/die Kollegen/ausbortchte.

105 mismatch Ganze/vierzig/Minuten/hatte/die/Nachbarin/das/Gerücht/erfahren/obwohl/die/Geheimhaltung/ihrer/Nachbarn.
correction In/vierzig/Minuten/hatte/die/Nachbarin/das/Gerücht/erfahren/obwohl/die/Geheimhaltung/ihrer/Nachbarn.
control Vor/vierzig/Minuten/hatte/die/Nachbarin/das/Gerücht/erfahren/obwohl/die/Geheimhaltung/ihrer/Nachbarn.

106 mismatch Ganze/dreißig/Minuten/hatte/der/Manager/der/Auftrag/erhalten/obwohl/die/Inhaberin/billiger/war.
correction In/dreißig/Minuten/hatte/der/Manager/der/Auftrag/erhalten/obwohl/die/Inhaberin/billiger/war.
control Vor/dreißig/Minuten/hatte/der/Manager/der/Auftrag/erhalten/obwohl/die/Inhaberin/billiger/war.

107 mismatch Ganze/drei/Stunden/hatte/der/Student/Innen/Platz/erhalten/trotz/der/überaus/langen/Warteschlange.
correction In/drei/Stunden/hatte/der/Student/Innen/Platz/erhalten/trotz/der/überaus/langen/Warteschlange.
control Vor/drei/Stunden/hatte/der/Student/Innen/Platz/erhalten/trotz/der/überaus/langen/Warteschlange.

correction In/zwei/Tagen/hatte/das/Mädchen/eine/Karte/erhalten/obwohl/der/Schwarzmarkt/quasi/leergefegt/war.
control Vor/zwei/Tagen/hatte/das/Mädchen/eine/Karte/erhalten/obwohl/der/Schwarzmarkt/quasi/leergefegt/war.


correction In/dreißig/Minuten/hatte/die/Polizei/den/Einbrecher/ertapppt/obwohl/dieser/alle/Eventualitäten/abschätzte.
control Vor/dreißig/Minuten/hatte/die/Polizei/den/Einbrecher/ertapppt/obwohl/dieser/alle/Eventualitäten/abschätzte.
The Processing of Events

### Table K.2. "Tense"-Items

<table>
<thead>
<tr>
<th>item</th>
<th>condition</th>
<th>sentence</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>violation control</td>
<td>Übermorgen/um acht/kam/die/alte/Frau/in/das/Cafe. Vorgestern/um acht/kam/die/alte/Frau/in/das/Cafe.</td>
</tr>
</tbody>
</table>

2nd proofs
<table>
<thead>
<tr>
<th>item</th>
<th>condition</th>
<th>sentence</th>
</tr>
</thead>
<tbody>
<tr>
<td>19</td>
<td>violation</td>
<td>Morgen/Vormittag/skizzierte/der/ehrgeizige/Manager/seine/aktuellen/Pläne. Gestern/Vormittag/skizzierte/der/ehrgeizige/Manager/seine/aktuellen/Pläne.</td>
</tr>
</tbody>
</table>
Letzte/Woche/strickte/die/nette/Großmutter/mehrere/Paar/Socken.

Nächsten/Woche/ehrte/die/kleine/Gemeinde/den/berühmten/Dichter.

Übermorgen/Abend/mixte/der/Barmixer/Longdrinks/für/die/Gäste.

Gestern/Abend/küsste/der/verliebte/Junge/das/schüchterne/Mädchen.

Morgen/Nacht/wuschte/der/Barkeeper/ins/Trenen/mit/einem/Tisch.

Nächsten/Montag/montierte/der/Monteur/den/Motor/der/Maschine.

Übernächste/Woche/programmierte/der/Informatiker/das/längst/überfällige/Programm.

Übernächste/Woche/programmierte/der/Informatiker/das/längst/überfällige/Programm.

Letzte/Woche/strickte/die/nette/Großmutter/mehrere/Paar/Socken.

Übermorgen/Mittag/entfernte/der/englische/Gärtner/das/wuchernde/Unkraut.

Vorgestern/Abend/brach/die/Expedition/in/die/Wüste/auf.

Kommendes/Jahr/erforschte/der/Professor/das/Wachstum/dieser/Krebzellen.

Übernächste/Woche/programmierte/der/Informatiker/das/längst/überfällige/Programm.

Letzte/Woche/strickte/die/nette/Großmutter/mehrere/Paar/Socken.

Letzte/Woche/arbeitete/das/eingespielte/Team/am/neuen/Auftrag.

Übernächste/Woche/programmierte/der/Informatiker/das/längst/überfällige/Programm.
### APPENDIX L

**Normed fillers**

<table>
<thead>
<tr>
<th>cat.</th>
<th>sentence</th>
<th>judgment</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>Er glaubt, dass der Lehrer den Schüler zu Unrecht ausgeschimpft hat.</td>
<td>.93</td>
</tr>
<tr>
<td>A2</td>
<td>Das freche Mädchen hat es gewagt, ihrem Bruder ein Stück Seife zu schenken.</td>
<td>.94</td>
</tr>
<tr>
<td>A3</td>
<td>Der Patient hat den geldgierigen Zahnarzt überlistet.</td>
<td>1.03</td>
</tr>
<tr>
<td>A4</td>
<td>Der Stürmer hat den Torwart vorsätzlich gefoult.</td>
<td>1.13</td>
</tr>
<tr>
<td>B1</td>
<td>Er glaubt, der Komponist hat den Solisten in dieser Arie überfordert.</td>
<td>.56</td>
</tr>
<tr>
<td>B2</td>
<td>Sie hofft, das Finanzamt hat den Betrüger überlistet.</td>
<td>.28</td>
</tr>
<tr>
<td>B3</td>
<td>Der Kaiser hat dem Fürsten den Maler empfohlen.</td>
<td>.42</td>
</tr>
<tr>
<td>B4</td>
<td>Er fürchtet, der Soldat hat den Demonstranten absichtlich angeschossen.</td>
<td>.27</td>
</tr>
<tr>
<td>C1</td>
<td>In Rottenburg fürchtet er, hat der Händler den Politiker bestochen.</td>
<td>–.49</td>
</tr>
<tr>
<td>C2</td>
<td>Der Bischof hat ihm es anvertraut.</td>
<td>–.52</td>
</tr>
<tr>
<td>C3</td>
<td>Die Frau, die die Katze, die die Maus fing, fütterte, seufzte.</td>
<td>–.15</td>
</tr>
<tr>
<td>C4</td>
<td>Der Kardinal fürchtet er, hat den Theologen beleidigt.</td>
<td>–1.00</td>
</tr>
<tr>
<td>D1</td>
<td>Der Komponist hat dem neuen Tenor es zugemutet.</td>
<td>–1.01</td>
</tr>
<tr>
<td>D2</td>
<td>Das Au-pair-Mädchen lernte dem Jungen jeden Morgen etwas Französisch.</td>
<td>–1.26</td>
</tr>
<tr>
<td>D3</td>
<td>Schade, dass dieses Buch alle ohne zu lesen ins Regal gestellt haben.</td>
<td>–.81</td>
</tr>
<tr>
<td>D4</td>
<td>Wir lesen am liebsten die Süddeutsche, obwohl wir leben jetzt in Düsseldorf.</td>
<td>–1.09</td>
</tr>
<tr>
<td>E1</td>
<td>Beim Stammtisch die drei Freunde spielen mit Vorliebe Skat.</td>
<td>–1.40</td>
</tr>
<tr>
<td>E2</td>
<td>Gelesen denke ich, dass das Buch keiner hat.</td>
<td>–1.72</td>
</tr>
<tr>
<td>E3</td>
<td>Der Waffenhändler glaubt er, dass den Politiker bestochen hat.</td>
<td>–1.77</td>
</tr>
<tr>
<td>E4</td>
<td>Der Dirigent bedauert er, dass den Sänger überfordert hat.</td>
<td>–1.78</td>
</tr>
</tbody>
</table>

*note: judgment values are mean z-scores, cat. = category.*
Table M.1. Items used in Experiment 10a

<table>
<thead>
<tr>
<th>item</th>
<th>condition</th>
<th>sentence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>short-for</td>
<td>Der Arbeiter/die Schubkarre/fünf Minuten lang./ dann/wurde er/ plötzlich/ woanders hin gerufen.</td>
</tr>
<tr>
<td></td>
<td>short-in</td>
<td>Der Arbeiter/die Schubkarre/in fünf Minuten./ dann/wurde er/ plötzlich/ woanders hin gerufen.</td>
</tr>
<tr>
<td></td>
<td>long-for</td>
<td>Der Arbeiter/die Schubkarre/fünf Jahre lang./ dann/wurde er/ endlich/ für einen neuen Job eingeteilt.</td>
</tr>
<tr>
<td></td>
<td>long-in</td>
<td>Der Arbeiter/die Schubkarre/in fünf Jahren./ dann/wurde er/ endlich/ für einen neuen Job eingeteilt.</td>
</tr>
<tr>
<td>2</td>
<td>short-for</td>
<td>Der Athlet/lief/den Marathon/drei Stunden lang./ dann/wurde er/ erschöpft/ von der Bahn getragen.</td>
</tr>
<tr>
<td></td>
<td>short-in</td>
<td>Der Athlet/lief/den Marathon/in drei Stunden./ dann/wurde er/ erschöpft/ von der Bahn getragen.</td>
</tr>
<tr>
<td></td>
<td>long-for</td>
<td>Der Athlet/lief/den Marathon/zwanzig Jahre lang./ dann/wurde er/ langsam zu alt/ dafür.</td>
</tr>
<tr>
<td></td>
<td>long-in</td>
<td>Der Athlet/lief/den Marathon/in zwanzig Jahren./ dann/wurde er/ langsam zu alt/ dafür.</td>
</tr>
<tr>
<td>3</td>
<td>short-for</td>
<td>Der Journalist/verfasste/die Kolumne/zwanzig Minuten lang./ dann/wurde er/ plötzlich/ zum Chefredakteur gerufen.</td>
</tr>
<tr>
<td></td>
<td>short-in</td>
<td>Der Journalist/verfasste/die Kolumne/in zwanzig Minuten./ dann/wurde er/ plötzlich/ zum Chefredakteur gerufen.</td>
</tr>
<tr>
<td></td>
<td>long-for</td>
<td>Der Journalist/verfasste/die Kolumne/zwanzig Monate lang./ dann/wurde er/ überraschend/ zum Chefredakteur befördert.</td>
</tr>
<tr>
<td></td>
<td>long-in</td>
<td>Der Journalist/verfasste/die Kolumne/in zwanzig Monaten./ dann/wurde er/ überraschend/ zum Chefredakteur befördert.</td>
</tr>
<tr>
<td>4</td>
<td>short-for</td>
<td>Die Köchin/kochte/das Tagesessen/eine Stunde lang./ dann/wurde sie/ von einem Herrn vom WKD überrascht.</td>
</tr>
<tr>
<td></td>
<td>short-in</td>
<td>Die Köchin/kochte/das Tagesessen/in einer Stunde./ dann/wurde sie/ von einem Herrn vom WKD überrascht.</td>
</tr>
<tr>
<td></td>
<td>long-for</td>
<td>Die Köchin/kochte/das Tagesessen/ein Jahr lang./ dann/wurde sie/ plötzlich/ ohne weitere Begründung entlassen.</td>
</tr>
<tr>
<td></td>
<td>long-in</td>
<td>Die Köchin/kochte/das Tagesessen/in einem Jahr./ dann/wurde sie/ plötzlich/ ohne weitere Begründung entlassen.</td>
</tr>
</tbody>
</table>
     long-for  Die Mutter/reinigte/das Kinderzimmer/zwei Jahre lang,/dann/ wurde sie/endlich/wütend/ auf ihre faule Tochter.
     long-in  Die Mutter/reinigte/das Kinderzimmer/in zwei Jahren,/dann/ wurde sie/endlich/wütend/ auf ihre faule Tochter.

6  short-for  Helga/putzte/das Bad/einen Monat lang,/dann/hatte sie/endlich/
     short-in  Helga/putzte/das Bad/in einem Monat,/dann/hatte sie/endlich/
     long-for  Helga/putzte/das Bad/zweie Jahre lang,/dann/
     long-in  Helga/putzte/das Bad/zweie Jahren,

7  short-for  Franz/staubzte/das Wohnzimmer/in zwanzig Minuten,/dann/
     short-in  Franz/staubzte/das Wohnzimmer/in zwanzig Minuten,/dann/
     long-for  Franz/staubzte/das Wohnzimmer/in zwanzig Minuten,/dann/
     long-in  Franz/staubzte/das Wohnzimmer/in zwanzig Minuten,/dann/

8  short-for  Der Chauffeur/durchfuhr/den Tunnel/in fünf Minuten,/dann/erst/ wurde er/auf die blinkende Öllampe/ aufmerksam.
     short-in  Der Chauffeur/durchfuhr/den Tunnel/in fünf Minuten,/dann/erst/ wurde er/auf die blinkende Öllampe/ aufmerksam.
     long-for  Der Chauffeur/durchfuhr/den Tunnel/in fünf Minuten,/dann/erst/ wurde er/auf die blinkende Öllampe/ aufmerksam.
     long-in  Der Chauffeur/durchfuhr/den Tunnel/in fünf Minuten,/dann/erst/ wurde er/auf die blinkende Öllampe/ aufmerksam.

9  short-for  Der kleine Hans/spielte/die Sonate/in drei Wochen,/dann/ hatte er/keine Lust mehr/auf das Stück.
     short-in  Der kleine Hans/spielte/die Sonate/in drei Wochen,/dann/ hatte er/keine Lust mehr/auf das Stück.
     long-for  Der kleine Hans/spielte/die Sonate/in drei Wochen,/dann/ hatte er/keine Lust mehr/auf das Stück.
     long-in  Der kleine Hans/spielte/die Sonate/in drei Wochen,/dann/ hatte er/keine Lust mehr/auf das Stück.

10 short-for  Der Hobbypilot/überquerte/die Alpen/zwei Stunden lang,/dann/ wurde er/durch einen Sturm/zur Landung gezwungen.
     short-in  Der Hobbypilot/überquerte/die Alpen/zwei Stunden lang,/dann/ wurde er/durch einen Sturm/zur Landung gezwungen.
     long-for  Der Hobbypilot/überquerte/die Alpen/zwei Stunden lang,/dann/ wurde er/durch einen Sturm/zur Landung gezwungen.
     long-in  Der Hobbypilot/überquerte/die Alpen/zwei Stunden lang,/dann/ wurde er/durch einen Sturm/zur Landung gezwungen.
<table>
<thead>
<tr>
<th>item</th>
<th>condition</th>
<th>sentence</th>
</tr>
</thead>
<tbody>
<tr>
<td>16</td>
<td>short-for</td>
<td>Der Mieter/reparierte/den Wasserhahn/eine Stunde lang,/dann erst/ hatte er/sich beim Vermieter/beschwert.</td>
</tr>
<tr>
<td></td>
<td>short-in</td>
<td>Der Mieter/reparierte/den Wasserhahn/in einer Stunde,/dann erst/ hatte er/sich beim Vermieter/beschwert.</td>
</tr>
<tr>
<td></td>
<td>long-for</td>
<td>Der Mieter/reparierte/den Wasserhahn/fünf Jahre lang,/dann erst/ hatte er/vom Vermieter/einen neuen bekommen.</td>
</tr>
<tr>
<td></td>
<td>long-in</td>
<td>Der Mieter/reparierte/den Wasserhahn/in fünf Jahren,/dann erst/ hatte er/vom Vermieter/einen neuen bekommen.</td>
</tr>
<tr>
<td>17</td>
<td>short-for</td>
<td>Herr Maier/überquerte/die Bundesstrasse/zwanzig Sekunden lang,/ dann/wurde er/von einem Polizisten/angehalten.</td>
</tr>
<tr>
<td></td>
<td>short-in</td>
<td>Herr Maier/überquerte/die Bundesstrasse/in zwanzig Sekunden,/ dann/wurde er/von einem Polizisten/angehalten.</td>
</tr>
<tr>
<td></td>
<td>long-for</td>
<td>Herr Maier/überquerte/die Bundesstrasse/zwanzig Tage lang,/dann/ wurde er/auf die neue Brücke/ aufmerksam.</td>
</tr>
<tr>
<td></td>
<td>long-in</td>
<td>Herr Maier/überquerte/die Bundesstrasse/in zwanzig Tagen,/dann/ wurde er/auf die neue Brücke/ aufmerksam.</td>
</tr>
<tr>
<td>18</td>
<td>short-for</td>
<td>Der Buchhalter/verfasste/die Monatsbilanz/drei Wochen lang,/ dann/wurde er/plötzlich/entlassen.</td>
</tr>
<tr>
<td></td>
<td>short-in</td>
<td>Der Buchhalter/verfasste/die Monatsbilanz/in drei Wochen,/dann/ wurde er/plötzlich/entlassen.</td>
</tr>
<tr>
<td></td>
<td>long-for</td>
<td>Der Buchhalter/verfasste/die Monatsbilanz/sechs Jahre lang,/dann/ wurde er/plötzlich/entlassen.</td>
</tr>
<tr>
<td></td>
<td>long-in</td>
<td>Der Buchhalter/verfasste/die Monatsbilanz/in sieben Jahren,/dann/ wurde er/plötzlich/entlassen.</td>
</tr>
<tr>
<td></td>
<td>short-in</td>
<td>Das Orchester/schrieb/die Ouvertüre/in drei Minuten,/dann/ wurde es/durch Lärm im Publikum/gestört.</td>
</tr>
<tr>
<td></td>
<td>long-for</td>
<td>Das Orchester/schrieb/die Ouvertüre/drei Monate lang,/dann/ wurde es/den Musikern/aberdauernd zu viel.</td>
</tr>
<tr>
<td></td>
<td>long-in</td>
<td>Das Orchester/schrieb/die Ouvertüre/in drei Monaten,/dann/ wurde es/den Musikern/aberdauernd zu viel.</td>
</tr>
<tr>
<td></td>
<td>short-in</td>
<td>Der Sprecher/verlas/die Nachrichten/in zwei Minuten,/dann/ wurde er/von einem Schwindelanfall/überwältigt.</td>
</tr>
<tr>
<td>item</td>
<td>condition</td>
<td>sentence</td>
</tr>
<tr>
<td>------</td>
<td>-----------</td>
<td>----------</td>
</tr>
<tr>
<td></td>
<td>short-in</td>
<td>Die Sportlerin/durchschwamm/den Armelkanal/in zwölf Stunden,/dann/wurde sie/auf dem Sofa versprengt.</td>
</tr>
<tr>
<td></td>
<td>long-for</td>
<td>Die Sportlerin/durchschwamm/den Armelkanal/mehrere Jahre lang,/dann/wurde sie/langsam zu alt/zu dafür.</td>
</tr>
<tr>
<td></td>
<td>long-in</td>
<td>Die Sportlerin/durchschwamm/den Armelkanal/in mehreren Jahren,/dann/wurde sie/langsam zu alt/zu dafür.</td>
</tr>
<tr>
<td></td>
<td>short-in</td>
<td>Der Dieb/leerte/die Kasse/in einer Minute,/dann/wurde er/von der Polizei/festgenommen.</td>
</tr>
<tr>
<td></td>
<td>long-for</td>
<td>Der Dieb/leerte/die Kasse/zu einem Jahr lang,/dann/wurde er/von der Polizei/festgenommen.</td>
</tr>
<tr>
<td></td>
<td>long-in</td>
<td>Der Dieb/leerte/die Kasse/in einem Jahr,/dann/wurde er/von der Polizei/festgenommen.</td>
</tr>
<tr>
<td>29</td>
<td>short-for</td>
<td>Der Barbier/rasierte/den Peter/drei Minuten lang,/dann/wurde er/die Klinge/stumpf.</td>
</tr>
<tr>
<td></td>
<td>short-in</td>
<td>Der Barbier/rasierte/den Peter/in drei Minuten,/dann/wurde er/die Klinge/stumpf.</td>
</tr>
<tr>
<td></td>
<td>long-for</td>
<td>Der Barbier/rasierte/den Peter/drei Wochen lang,/dann/wurde er/die Klinge/stumpf.</td>
</tr>
<tr>
<td></td>
<td>long-in</td>
<td>Der Barbier/rasierte/den Peter/in drei Wochen,/dann/wurde er/die Klinge/stumpf.</td>
</tr>
<tr>
<td></td>
<td>long-in</td>
<td>Sabine/rezitierte/das Gedicht/in zwei Monaten,/dann/konnte sie/es/nicht mehr hören.</td>
</tr>
<tr>
<td>31</td>
<td>short-for</td>
<td>Der Student/al/das Mensaessen/zehn Minuten lang,/dann/musste er/schnell/zum nächsten Seminar laufen.</td>
</tr>
<tr>
<td></td>
<td>short-in</td>
<td>Der Student/al/das Mensaessen/in zehn Minuten,/dann/musste er/schnell/zum nächsten Seminar laufen.</td>
</tr>
<tr>
<td></td>
<td>long-for</td>
<td>Der Student/al/das Mensaessen/zehn Semester lang,/dann/musste er/nicht mehr dort/zu Mittag essen.</td>
</tr>
<tr>
<td></td>
<td>long-in</td>
<td>Der Student/al/das Mensaessen/in zehn Semestern,/dann/musste er/nicht mehr dort/zu Mittag essen.</td>
</tr>
</tbody>
</table>
### Appendix N

**Target sentences in Experiment 10b**

<table>
<thead>
<tr>
<th>item</th>
<th>condition</th>
<th>sentence</th>
</tr>
</thead>
<tbody>
<tr>
<td>38</td>
<td>short-for</td>
<td>Das Shuttle/durchflog/die Atmosphäre/zwanzig Minuten lang,/dann/war/Kurs auf den Weltraumhafen/gesetzt worden.</td>
</tr>
<tr>
<td></td>
<td>short-in</td>
<td>Das Shuttle/durchflog/die Atmosphäre/in zwanzig Minuten,/dann/war/Kurs auf den Weltraumhafen/gesetzt worden.</td>
</tr>
<tr>
<td></td>
<td>long-for</td>
<td>Das Shuttle/durchflog/die Atmosphäre/drei Jahre lang,/dann/war/die Raumstation/endlich fertig.</td>
</tr>
<tr>
<td></td>
<td>long-in</td>
<td>Das Shuttle/durchflog/die Atmosphäre/in drei Jahren,/dann/war/die Raumstation/endlich fertig.</td>
</tr>
<tr>
<td>39</td>
<td>short-for</td>
<td>Der Wildhüter/durchquerte/den Nationalpark/einen Tag lang,/dann/wurde er/per Handy/über den Unfall informiert.</td>
</tr>
<tr>
<td></td>
<td>short-in</td>
<td>Der Wildhüter/durchquerte/den Nationalpark/in einem Tag,/dann/wurde er/per Handy/über den Unfall informiert.</td>
</tr>
<tr>
<td></td>
<td>long-in</td>
<td>Der Wildhüter/durchquerte/den Nationalpark/in vielen Jahren,/dann/wurde er/von einem Grizzly/tödlich verletzt.</td>
</tr>
<tr>
<td>40</td>
<td>short-for</td>
<td>Der Briefträger/leerte/den Briefkasten/eine Minute lang,/dann/wurde er/auf den seltsamen Brief/aufmerksam.</td>
</tr>
<tr>
<td></td>
<td>short-in</td>
<td>Der Briefträger/leerte/den Briefkasten/in einer Minute,/dann/wurde er/auf den seltsamen Brief/aufmerksam.</td>
</tr>
<tr>
<td></td>
<td>long-for</td>
<td>Der Briefträger/leerte/den Briefkasten/zwei Jahre lang,/dann/wurde er/auf eine andere Tour/eingeteilt.</td>
</tr>
<tr>
<td></td>
<td>long-in</td>
<td>Der Briefträger/leerte/den Briefkasten/in zwei Jahren,/dann/wurde er/auf eine andere Tour/eingeteilt.</td>
</tr>
<tr>
<td>1</td>
<td>ago</td>
<td>Vor wenigen Minuten/nieste/die Schülerin/recht laut,/dann/ging sie/heim und/legte sich ins Bett.</td>
</tr>
<tr>
<td></td>
<td>for</td>
<td>Den ganzen Morgen/nieste/die Schülerin/recht laut,/dann/ging sie/heim und/legte sich ins Bett.</td>
</tr>
<tr>
<td>2</td>
<td>ago</td>
<td>Vor fünfzehn Minuten/knallte/ein Gewehr/beunruhigend nahe,/dann/wurde es/wieder still/um das Dorf.</td>
</tr>
<tr>
<td></td>
<td>for</td>
<td>Mehrere Tage lang/knallte/ein Gewehr/beunruhigend nahe,/dann/wurde es/wieder still/um das Dorf.</td>
</tr>
<tr>
<td>3</td>
<td>ago</td>
<td>Gerade eben/gähnte/der kleine Junge/vor dem Fernseher,/da es/bereits/sehr spät/geworden war.</td>
</tr>
<tr>
<td>4</td>
<td>ago</td>
<td>Vor fünf Minuten/donnerte/die Kanone/beim Festumzug,/als/die Faschingsgesellschaft/durch die Stadt/zog.</td>
</tr>
<tr>
<td></td>
<td>for</td>
<td>Fünf Minuten lang/donnerte/die Kanone/beim Festumzug,/als/die Faschingsgesellschaft/durch die Stadt/zog.</td>
</tr>
<tr>
<td>5</td>
<td>ago</td>
<td>Gerade eben/hustete/Peter/ziemlich laut,/um/Friedas Aufmerksamkeit/auf sich/zu ziehen.</td>
</tr>
<tr>
<td></td>
<td>for</td>
<td>Fünf Minuten lang/hustete/Peter/ziemlich laut,/um/Friedas Aufmerksamkeit/auf sich/zu ziehen.</td>
</tr>
<tr>
<td>6</td>
<td>ago</td>
<td>Vor wenigen Minuten/stolperte/der Rentner/in seiner Wohnung,/dann/kontaktierte/er/den Arzt.</td>
</tr>
<tr>
<td></td>
<td>for</td>
<td>Eine Woche lang/stolperte/der Rentner/in seiner Wohnung,/dann/kontaktierte/er/den Arzt.</td>
</tr>
<tr>
<td>7</td>
<td>ago</td>
<td>Vor wenigen Minuten/queikte/die kleine Maus/in ihrem Bau,/da/lieh sich/Frau Maier/die Nachbarskatze.</td>
</tr>
<tr>
<td></td>
<td>for</td>
<td>Mehrere Wochen/queikte/die kleine Maus/in ihrem Bau,/da/lieh sich/Frau Maier/die Nachbarskatze.</td>
</tr>
<tr>
<td>8</td>
<td>ago</td>
<td>Gerade eben/blickte/der besorgte Vater/aus dem kleinen Fenster,/da/entdeckte er/seine angebrunkene/Tochter.</td>
</tr>
<tr>
<td></td>
<td>for</td>
<td>Viele Stunden/blickte/der besorgte Vater/aus dem kleinen Fenster,/da/entdeckte er/seine angebrunkene/Tochter.</td>
</tr>
</tbody>
</table>
## Appendix N. Target sentences in Experiment 10b

### 2nd proofs

<table>
<thead>
<tr>
<th>item</th>
<th>condition</th>
<th>sentence</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>ago</td>
<td>Vor zehn Minuten/rief/der kleine Junge/um Hilfe,/dann/endlich/ wurde ein Feuerwehrmann/auf ihn aufmerksam.</td>
</tr>
<tr>
<td></td>
<td>for</td>
<td>Zehn Minuten lang/rief/der kleine Junge/um Hilfe,/dann/endlich/ wurde ein Feuerwehrmann/auf ihn aufmerksam.</td>
</tr>
<tr>
<td>10</td>
<td>ago</td>
<td>Gerade eben/piepste/der Computer/überaus laut,/dann/schmorte/ das Motherboard/zusammen.</td>
</tr>
<tr>
<td></td>
<td>for</td>
<td>Zehn Minuten lang/piepste/der Computer/überaus laut,/dann/ schmorte/das Motherboard/zusammen.</td>
</tr>
<tr>
<td>12</td>
<td>ago</td>
<td>Vor wenigen Minuten/ertönte/der Ruf/der Nachtigall,/dann/ war es/aber wieder/still im Wald.</td>
</tr>
<tr>
<td></td>
<td>for</td>
<td>Eine Stunde lang/ertönte/der Ruf/der Nachtigall,/dann/war es/ aber wieder/still im Wald.</td>
</tr>
<tr>
<td>14</td>
<td>ago</td>
<td>Vor ein paar Minuten/pfiff/das Mädchen/überaus laut,/dann/ kam ihre Freundin/auf die Strasse/zum Spielen.</td>
</tr>
<tr>
<td></td>
<td>for</td>
<td>Ein paar Minuten lang/pfiff/das Mädchen/überaus laut,/dann/ kam ihre Freundin/auf die Strasse/zum Spielen.</td>
</tr>
<tr>
<td>15</td>
<td>ago</td>
<td>Vor einer Viertelstunde/seufzte/die Patientin/vernehmlich,/dann/ fiel sie/wieder zurück/in die Bewußtlosigkeit.</td>
</tr>
<tr>
<td></td>
<td>for</td>
<td>Eine Viertelstunde lang/seufzte/die Patientin/vernehmlich,/dann/ fiel sie/wieder zurück/in die Bewußtlosigkeit.</td>
</tr>
<tr>
<td>16</td>
<td>ago</td>
<td>Heute Morgen/erzitterte/der Baum/im Wind,/dann/hatte sich/ der Sturm/aber rasch wieder gelegt.</td>
</tr>
<tr>
<td></td>
<td>for</td>
<td>Den ganzen Morgen/erzitterte/der Baum/im Wind,/dann/ hatte sich/der Sturm/aber rasch wieder gelegt.</td>
</tr>
<tr>
<td>17</td>
<td>ago</td>
<td>Vor wenigen Minuten/schaute/der Vater/in den Topf,/dann/deckte/ er/den Tisch.</td>
</tr>
<tr>
<td>18</td>
<td>ago</td>
<td>Gerade eben/räusperte sich/der Zuschauer/laut vernehmlich,/ dann/verließ/er/die Vorstellung.</td>
</tr>
<tr>
<td></td>
<td>for</td>
<td>Die ganze Zeit/räusperte sich/der Zuschauer/laut vernehmlich,/ dann/verließ/er/die Vorstellung</td>
</tr>
<tr>
<td>19</td>
<td>ago</td>
<td>Gerade eben/hüpfte/Martin/über die Mauer,/dann/verscheuchte ihn/der Nachbar/aus dem Garten.</td>
</tr>
<tr>
<td></td>
<td>for</td>
<td>Den ganzen Nachmittag/hüpfte/Martin/über die Mauer,/dann/ verscheuchte ihn/der Nachbar/aus dem Garten.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>item</th>
<th>condition</th>
<th>sentence</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>ago</td>
<td>Vor wenigen Minuten/ertönte/der Schlachtruf/im Tal,/dann/ zogen/die Reiter/in den Kampf.</td>
</tr>
<tr>
<td></td>
<td>for</td>
<td>Den ganzen Tag/ertönte/der Schlachtruf/im Tal,/dann/zogen/ die Reiter/in den Kampf.</td>
</tr>
</tbody>
</table>