SENTENCE COMPREHENSION OF EVENT STRUCTURE
IN ENGLISH AND JAPANESE:
AN EVALUATION OF THE INTERACTION
BETWEEN GRAMMATICAL ASPECT AND LEXICAL ASPECT

A DISSERTATION SUBMITTED TO THE GRADUATE DIVISION OF THE
UNIVERSITY OF HAWAI‘I AT MĀNOA IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
IN
LINGUISTICS

AUGUST 2014

By

Yukie Hara

Dissertation Committee:

Amy J. Schafer, Chairperson
  William O’Grady
  Kamil Ud Deen
  Shinichiro Fukuda
  Graham V. Crookes

Keywords: Time, Event, Sentence comprehension, Japanese, English
ACKNOWLEDGEMENTS

This dissertation reflects considerable assistance and friendship from many people that I am grateful to know in Hawai‘i and Japan. I treasure these friends, colleagues, teachers, and advisers who have encouraged me throughout the years of my academic journey. I would like to extend my aloha to all of them, and my particular gratitude to those mentioned here.

I would like to express my very great appreciation to my dissertation committee members for their thoughtful support and advice, which has enabled me to come this far. My more than great gratitude and appreciation goes to my adviser and dissertation committee chair, Amy J. Schafer, for her significant support over the years. It would NEVER have been possible for me to achieve my goals without her excellent advising, professional attitude, thoughtful concern, and vital encouragement all along. She has been the best possible guide, understanding exactly what I need and always nourishing and improving my work. Her kind concern and her constant help brought me through difficult periods. As a researcher, and as a person, I have learned a lot from her, and I feel very fortunate and honored to be her advisee. I am grateful for her extended support after I left Hawai‘i to continue my life while writing the dissertation in Japan. Without her, this dissertation would not have seen its completion. I thank her from the bottom of my heart. The constructive advice and encouragement given by William O’Grady has been greatly appreciated for making me think critically and deepening this dissertation’s treatment of the topic. I am grateful to Kamil Ud Deen for the kind support and useful criticism he has offered, and also for my meaningful experience as a research assistant on his language acquisition project in Thailand. I extend my deep gratitude to Shinichiro Fukuda, who has provided me with elaborate feedback; his encouraging support for my interest in Japanese syntax and semantics maximized my joy in working on the dissertation. I would like to express my sincere gratitude to Graham V. Crookes for his warm support and smiles. My special thanks go to Yuko Otsuka, as a proxy, and to John Haig, now retired, for taking part in my dissertation committee at earlier times.

The data collection would not have been possible without the tremendous help offered by many people. I would like to extend great thanks to the students at the University of Hawai‘i, Tohoku University (Miyagi, Japan), and Osaka University (Osaka, Japan), for their participation in my experiments. I enjoyed working with them, and I appreciated their willingness to share their time, their positive attitudes, and their lively interest in research with me. My gratitude also goes to the librarians at the two universities in Japan for offering me the resources and spaces for running experiments. The thoughtful and
crucial support for the data collection given by my former advisers at Tohoku University and by Toshiaki Furukawa, then of Osaka University, is deeply appreciated.

I would like to express my sincere gratitude to my former advisers at Tohoku University (Sendai, Japan): Naoyuki Ono, Takeshi Nakamoto, Tadao Miyamoto, Heiko Narrog, Yoshio Kawahira, and Daiko Takahashi. I owe them many thanks for their continuous support in the days before (and after) I came to the University of Hawai‘i. In particular, I cannot thank Naoyuki Ono enough for his thoughtful understanding and support, and for cultivating my research interest in aspect studies. His warm understanding and support of my decision to study abroad, without which I would not have had such a rich academic journey, earned my lasting gratitude. I would also like to extend my gratitude to Masaru Nakamura and Yosuke Sato for inspiring me to proceed deeper into the wonderful world of linguistics.

I wish to extend a grateful acknowledgement of the support and deep understanding of language I found among the faculty members at the Department of Linguistics and related departments at the University of Hawai‘i. My special thanks go to Yasuhiro Shirai at the University of Pittsburgh for the enlightening and encouraging communications regarding aspect studies.

This dissertation would not have been completed without financial support from various sources. I would like to express my very great appreciation to the Crown Prince Akihito Scholarship Foundation for supporting me during my first two years in Hawai‘i, making my otherwise-given-up-dream-of-studying-abroad come true. My appreciation goes to the College of Arts and Sciences and the Department of Linguistics at the University of Hawai‘i for grants funding my dissertation research, and also to the Department of Linguistics for the opportunity to work in graduate assistantships during my four years in Hawai‘i. I also wish to acknowledge Ochanomizu University (Tokyo, Japan) for making it possible for me to continue writing my dissertation while working for them in my transition from Hawai‘i to Japan.

I have been greatly helped by proofreaders; I wish to extend my great gratitude and acknowledgement for the continuous and significant kind help provided by Laurie Durand. My gratitude also goes to Tobias Bloyd and Stephanie Tunçay for helping prepare my experimental sentences in another language. I would like to acknowledge the excellent assistance provided by the secretaries at the Department of Linguistics, Jennifer Kanda and Nora Lum. The technical advice I received from the late Kurt Brunner in running experiments was greatly appreciated.
I have spent wonderful years in Hawai‘i thanks to the people I met there, at the Department of Linguistics and elsewhere. I would like to thank them for having shared friendship with me and giving advice when I was in need; they have been a great help in pursuing this journey. I am particularly grateful to Manami Sato for her invaluable advice on psycholinguistics, longtime support, and great memories of time spent together. For their support in every respect and great times shared, I thank the Japanese students who are my seniors in the department, Kaori Ueki, Yumiko Enyo, Jun Nomura, and Toshiaki Furukawa, who really made my life in Hawai‘i feel like home. I would like to offer my many warm thanks to Jung-Hee Kim, Heeyeon Dennison (Yoon), Hyunah Ahn, Sunju Kim, Hiroko Sato, Nozomi Tanaka, Mari Miyao, Soyoung Kim, On-Soon Lee, Jin Sun Choe, Chae-Eun Kim, Kyae-Sung Park, Karen Huang, Nian Liu, Suann Robinson, Mayumi Oiwa, Apay (Ai-Yu) Tang, Hunter Hatfield, Hye-Young Kwak, Aya Takeda, Carl Polley, Jawee Perla, Jake Terrell, Bodo Winter, Tsai-Hsiu (Dorinda) Liu, Sang-Gu Kang, Jinsook Kim, Kum-Jeong Joo, Jonny Kim, Elaine Lau, Antonio Cheung, John Van Way, Diana Stojanovic, Aya Inoue, Satoko Kimura, Mai Takemoto, Megumi Gregorio, Kevin Chang, and the late Dan X. Hall. My appreciation for their friendship and the international experiences we have shared also goes to members of the East-West Center and the staff and scholars at the Crown Prince Akihito Scholarship Foundation. I would particularly like to acknowledge the support provided by Marsha Yokomichi, Vice President of the Crown Prince Akihito Scholarship Foundation, from my very first day in Hawai‘i when she warmly welcomed me, and through all the following years of friendship and hospitality.

I wish to extend my special gratitude to Masatoshi Takeda for his considerate support and essential encouragement during the time after I left Hawai‘i to continue writing the dissertation in Japan. I would like to express my very great gratitude and appreciation to my family, Kesami Kodaira and Michiko and Yasunori Hara, who have always been supportive of me, who believed that I could achieve my goals, and who encouraged me throughout these years, sharing my long journey, with their love and joy.
ABSTRACT

When we describe situations in daily life, we use temporal cues in language, such as tense, properties of events in verbs (lexical aspect), and forms of progressive or perfective (grammatical aspect). These cues play a role in conveying and interpreting when an event takes place and whether the event is ongoing or completed. Psycholinguistic studies have reported effects of these cues in sentence processing, although the possibility of interaction effects between them has been underexplored. Attested effects of grammatical aspect may be those of interaction with lexical aspect. Certain combinations of lexical and grammatical aspect are favored in child language across languages, and recent studies have reported that certain combinations also facilitate adult sentence processing. Moreover, the effect of grammatical aspect may not equally influence an entire event representation, but focus on specific event parts highlighted by interaction with lexical aspect.

This dissertation focuses on interaction between grammatical aspect and other sources of temporal cues in sentence processing. Experiment 1 examined English processing in cases of mismatch between lexical and grammatical aspect. Experiments 2, 3, 4, and 5 tested English and Japanese to identify whether the higher processing cost of imperfective over perfective sentences comes from event durativity or grammatical aspect. Event focus is explored in Japanese, where the imperfective aspect can focus on ongoing event parts (progressive reading) or an end-state (resultative reading), by interaction with lexical aspect. Experiments 6, 7, 8, and 9 investigated the resiliency of
event representations marked by grammatical aspect in light of the role of event focus in Japanese.

The findings did not support the study’s original assumptions that the interaction of grammatical aspect and lexical aspect has an effect on sentence processing. Grammatical aspect’s manifestation of its role and effects is not interfered with by interaction with lexical aspect; at the same time, its psychological reality in human sentence processing is not as universally stable as previous studies assume.
# TABLE OF CONTENTS

Acknowledgments...........................................................................................................iii
Abstract...............................................................................................................................vi
List of Tables......................................................................................................................xii
List of Figures....................................................................................................................xiii

Chapter 1: Introduction........................................................................................................1

Chapter 2: Grammatical aspect and lexical aspect in human language.........................10
  2.1 Time and event in language....................................................................................10
  2.2 Lexical aspect in syntax and semantics...............................................................10
  2.3 Grammatical aspect in relation to lexical aspect..................................................13
  2.4 Aspect in child language......................................................................................16
  2.5 Aspect in child Japanese......................................................................................26

Chapter 3: Grammatical aspect and lexical aspect in sentence processing:
  roles and interaction........................................................................................................30
  3.1 Situation model construction................................................................................30
      3.1.1 Situation model and grammatical aspect.....................................................31
  3.2 Psycholinguistic studies on lexical aspect............................................................40
  3.3 Exploring interaction between grammatical and lexical aspect............................43
      3.3.1 Interaction of lexical and grammatical aspect in mismatching
          combinations......................................................................................................44
      3.3.2 Interaction of lexical and grammatical aspect in sentence processing,
          and understudied potential effect from focusing change...............................47
      3.3.3 Interaction of lexical and grammatical aspect in resiliency of event
          representation in memory.................................................................................51

Chapter 4: Role of grammatical aspect with mismatched lexical aspect in construction of
  mental event representations..........................................................................................54
  4.1 Perfective grammatical aspect and lexical aspect without endpoint.......................55
      4.1.1 Participants.....................................................................................................55
      4.1.2 Materials.......................................................................................................56
      4.1.3 Procedure....................................................................................................60
      4.1.4 Conditions...................................................................................................61
      4.1.5 Predictions...................................................................................................61
4.1.6 Data analysis .................................................62
4.1.7 Results ......................................................65
4.1.8 Discussion ..................................................70

Chapter 5: Role of grammatical aspect and event information in on-line processing costs ........................................81

5.1 Event duration read after grammatical aspect (English) ......................87
  5.1.1 Participants ................................................87
  5.1.2 Materials ..................................................87
  5.1.3 Procedure ...............................................90
  5.1.4 Conditions ..............................................91
  5.1.5 Predictions .............................................92
  5.1.6 Data analysis ...........................................93
  5.1.7 Results ..................................................96
  5.1.8 Discussion ...............................................101

5.2 Event duration read before grammatical aspect (Japanese) ............105
  5.2.1 Participants (Progressive, Resultative) ..........................106
  5.2.2 Materials (Progressive, Resultative) ..........................106
  5.2.3 Procedure (Progressive, Resultative) ........................111
  5.2.4 Conditions (Progressive, Resultative) ........................112
  5.2.5 Predictions (Progressive, Resultative) .........................112
  5.2.6 Data analysis ............................................115
    5.2.6.1 Progressive (Experiment 3) ............................115
    5.2.6.2 Resultative (Experiment 4) ............................117
    5.2.6.3 Residual reading times (Experiments 3 and 4) .........118
  5.2.7 Results ..................................................119
    5.2.7.1 Progressive (Experiment 3) ............................119
    5.2.7.2 Resultative (Experiment 4) ............................122
  5.2.8 Discussion ..............................................124
    5.2.8.1 Progressive (Experiment 3) ............................124
    5.2.8.2 Resultative (Experiment 4) ............................129

5.3 Without event duration before grammatical aspect (Japanese) .......133
  5.3.1 Participants ..............................................133
  5.3.2 Materials ...............................................134
  5.3.3 Procedure ..............................................135
  5.3.4 Conditions ..............................................135
  5.3.5 Predictions ..............................................136
  5.3.6 Data analysis ............................................136
  5.3.7 Results ..................................................139
  5.3.8 Discussion ...............................................141
Chapter 6: Role of grammatical aspect and focus on event parts in memory retrieval...155

6.1 Memory retrieval in short span..............................................................159
   6.1.1 Participants (Progressive, Resultative) ......................................160
   6.1.2 Materials (Progressive, Resultative) ............................................160
   6.1.3 Procedure (Progressive, Resultative) .........................................165
   6.1.4 Conditions (Progressive, Resultative) .......................................166
   6.1.5 Predictions (Progressive, Resultative) .......................................166
   6.1.6 Data analysis..................................................................................168
      6.1.6.1 Progressive (Experiment 6) ...............................................168
      6.1.6.2 Resultative (Experiment 7) ..............................................169
   6.1.7 Results..........................................................................................171
      6.1.7.1 Progressive (Experiment 6) ...............................................171
      6.1.7.2 Resultative (Experiment 7) ..............................................172
   6.1.8 Discussion......................................................................................174

6.2 Memory retrieval in long span............................................................177
   6.2.1 Participants (Progressive, Resultative) ......................................177
   6.2.2 Materials (Progressive, Resultative) ............................................177
   6.2.3 Procedure (Progressive, Resultative) .........................................178
   6.2.4 Conditions (Progressive, Resultative) .......................................178
   6.2.5 Predictions (Progressive, Resultative) .......................................178
   6.2.6 Data analysis..................................................................................180
      6.2.6.1 Progressive (Experiment 8) ...............................................181
      6.2.6.2 Resultative (Experiment 9) ..............................................182
   6.2.7 Results..........................................................................................184
      6.2.7.1 Progressive (Experiment 8) ...............................................184
      6.2.7.2 Resultative (Experiment 9) ..............................................185
   6.2.8 Discussion......................................................................................187

6.3 General discussion...............................................................................189

Chapter 7: General discussion .................................................................199

   7.1 Original claims and new claim.......................................................199
   7.2 Implications and ramifications.......................................................201
   7.3 Conclusion.......................................................................................208

Appendix A: Experiment 1, Critical preambles and prompts in two-alternative forced choice (2AFC) .................................................................210
Appendix B: Sub-analyses in Experiment 1..............................................214
Appendix C: Experiments 2 (English), Critical sentences.................................218
Appendix D: Experiments 3 (Japanese, Progressive), Critical sentences...........220
Appendix E: Experiments 4 (Japanese, Resultative), Critical sentences..........222
Appendix F: Experiments 5 (Japanese, Control), Critical sentences..............225
Appendix G: Experiments 6 and 8 (Progressive), Critical sentences.............227
Appendix H: Experiments 7 and 9 (Resultative), Critical sentences...............236

References.............................................................................................................245
## LIST OF TABLES

| Table                                                                 | Page |
|                                                                     |      |
| 2.1. Verb classification                                             | 12   |
| 5.1. Structure of the critical sentences in Experiment 2             | 88   |
| 5.2. Raw (and residual) reading times per word in each region from Experiment 2 | 98   |
| 5.3. Structure of the critical sentences in Experiment 3 (Progressive) | 107  |
| 5.4. Structure of the critical sentences in Experiment 4 (Resultative) | 107  |
| 5.5. Raw (and residual) reading times per word in each region from Experiment 3 | 120  |
| 5.6. Raw (and residual) reading times per word in each region from Experiment 4 | 122  |
| 5.7. Structure of the critical sentences in Experiment 5             | 134  |
| 5.8. Raw (and residual) reading times per word in each region from Experiment 5 | 140  |
| 6.1. Structure of the critical stories                              | 161  |
| 6.2. Positions of critical probes for Experiments 6 and 7            | 163  |
| 6.3: Positions of critical probes for Experiments 8 and 9            | 178  |
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1. Mean source preference ratio after reading a Perfective-Singular, Perfective-Plural, or Imperfective-Singular preamble in Experiment 1</td>
<td>66</td>
</tr>
<tr>
<td>4.2. Mean decision times in selecting Source or Goal referent after reading Perfective-Singular, Perfective-Plural, or Imperfective-Singular preamble in Experiment 1</td>
<td>68</td>
</tr>
<tr>
<td>5.1. Mean residual reading times in each region from Experiment 2</td>
<td>97</td>
</tr>
<tr>
<td>5.2. Mean raw reading times in each region from Experiment 2</td>
<td>97</td>
</tr>
<tr>
<td>5.3. Mean raw reading times in each region from Experiment 3</td>
<td>120</td>
</tr>
<tr>
<td>5.4. Mean raw reading times in each region from Experiment 4</td>
<td>122</td>
</tr>
<tr>
<td>5.5. Mean raw reading times in each region in the filler items from Experiment 3</td>
<td>128</td>
</tr>
<tr>
<td>5.6. Mean raw reading times in each region in the filler items from Experiment 4</td>
<td>132</td>
</tr>
<tr>
<td>5.7. Mean raw reading times in each region from Experiment 5</td>
<td>140</td>
</tr>
<tr>
<td>5.8. Mean raw reading times in each region in the filler items from Experiment 5</td>
<td>144</td>
</tr>
<tr>
<td>5.9. Mean residual reading times in each region in the filler items from Experiment 5</td>
<td>144</td>
</tr>
<tr>
<td>6.1. Mean decision latencies at different word positions from Experiment 6</td>
<td>172</td>
</tr>
<tr>
<td>6.2. Mean decision latencies at different word positions from Experiment 7</td>
<td>173</td>
</tr>
<tr>
<td>6.3. Mean decision latencies at different word positions from Experiment 8</td>
<td>185</td>
</tr>
<tr>
<td>6.4. Mean decision latencies at different word positions from Experiment 9</td>
<td>186</td>
</tr>
<tr>
<td>B.1. Mean source preference ratio after reading a Perfective-Singular, Perfective-Plural, or Imperfective-Singular preamble with distinction of iterativeness in Experiment 1</td>
<td>215</td>
</tr>
<tr>
<td>B.2. Mean source preference ratio after reading a Perfective-Singular, Perfective-Plural, or Imperfective-Singular preamble with distinction of co-locatedness in Experiment 1</td>
<td>216</td>
</tr>
</tbody>
</table>
B.3. Mean source preference ratio after reading a Perfective-Singular, Perfective-Plural, or Imperfective-Singular preamble with distinction of Pronoun set vs. Noun set in Experiment 1.
CHAPTER 1

INTRODUCTION

Human language encodes temporal properties of events by means of lexical and grammatical aspect. A verb (phrase) denotes situation types by its lexical aspect, also known as *aktionsart*, which encodes internal temporal properties of events in terms of durativity, telicity, and dynamicity. Based on these event properties that define aspectual classes, verbs have been classified into four verb classes: *state, activity, achievement*, and *accomplishment* (Vendler, 1957). Grammatical aspect, or “viewpoint” aspect, on the other hand, imposes a viewpoint onto an event; imperfective takes an inside viewpoint to see an event as on-going, while perfective takes an outside viewpoint to see an event as completed. Across languages, the major contrast in grammatical aspect is imperfective versus perfective (Comrie, 1976). For example, progressive morphology is used to denote imperfective aspect and simple past morphology is used to denote perfective aspect in English.

However, event representation, unlike merely decoding text information, is a dynamic enterprise performed by a language comprehender in sentence processing (Johnson-Laird, 1983; van Dyk & Kintsch, 1983): comprehension of language involves construction of the state of affairs, or situations, that are described by a text. This process is representatively captured as *situation models construction* (Glenberg, Meyer, & Lindem, 1987; also see Zwaan & Radvansky, 1998). Researchers in psycholinguistics have detected that forms at various levels act as linguistic cues, contributing to such
online event representations (Fletcher & Chrysler, 1990; Glenberg et al., 1987; Morrow, Bower, & Greenspan, 1989; Morrow, Greenspan, & Bower, 1987). The effect of grammatical aspect has been intensively explored in the framework of situation model construction (e.g., Carreiras, Carriedo, Alonso, & Fernandez, 1997; Ferretti, Kutas, & McRae, 2007; Madden & Zwaan, 2003; Magliano & Schleich, 2000; Zwaan, Madden, & Whitten, 2000). Lexical aspect, on the other hand, encodes the core semantics for event interpretation in verb phrases. Lexical aspect has drawn some attention in psycholinguistic research, but for the most part in relation to its role in syntactic processing rather than in construction of situation models. I see an emerging importance in exploring these streams of research by placing more focus on the interaction of lexical and grammatical aspect, rather than on each one independently.

The importance of the interaction between lexical and grammatical aspect is highlighted by empirical evidence from the field of language acquisition. Briefly, unlike adults, children are attested to undergeneralize tense-aspect morphology to semantically well-defined subclasses of verbs (i.e., the aspect hypothesis). English-speaking children are known to initially use past/perfective morphology predominantly with telic verbs (i.e., achievement and accomplishment verbs), and imperfective morphology with atelic verbs (i.e., action verbs) under the age of approximately 2;6. Such a consistent distribution of verb morphology in relation to specific verb types has been cross-linguistically attested in children’s early language production in English (Bloom, Lifter, & Hafitz, 1980; Harner, 1981; Olsen & Weinberg, 1999; Shirai & Andersen, 1995), among other languages. Studies report similar patterns in Japanese (Rispoli, 1981; Shirai 1993), but the evidence is mixed (Shirai 1998).
This dissertation research explores the role of grammatical aspect through detailed examination of the understudied interaction between grammatical aspect and lexical aspect. With the aspect hypothesis as their motivation, some psycholinguistic studies, such as those conducted by Yap, Chu, Yiu, Wong, Kwan, Matthews, Tan, Li, & Shirai (2009) and Yap, Inoue, Shirai, Matthews, Wong, Chan (2006), have reported evidence of an advantage of a certain combination of grammatical and lexical aspect in adult language comprehension (i.e., it facilitates processing speed). Still unknown is the role of grammatical aspect in a “mismatching” combination in sentence processing, and therefore the question of whether grammatical aspect uniformly exerts its role across different verb types remains unanswered. A likely possibility is that such a combination may have an impact on processing procedures, even affecting how we interpret described situations. Therefore, I first examine the processing of “mismatching” combinations of lexical and grammatical aspects, which previous studies have neglected.

I then explore whether the focus that is interactively determined by a combination of lexical and grammatical aspects on a described event is of importance to the role of grammatical aspect in language comprehension. The preceding research has revealed that activities described with the imperfective have a higher probability of being considered ongoing than those described with the perfective (Magliano & Schleich, 2000), and that grammatical aspect, specifically the imperfective aspect in contrast to the perfective aspect, increases or maintains the accessibility of an event representation and of the world knowledge that is related to the events (Carreiras et al., 1997; Ferretti, Kutas, & McRae, 2007; Ferretti, Rohde, Kehler, & Crutchley, 2009; Magliano & Schleich, 2000; Madden & Therriault, 2009; Madden & Zwaan, 2003; Morrow, 1990; Truitt & Zwaan, 1997).
However, although playing such a role, grammatical aspect may not exert its effect equally on the entire event in a representation; the effect may depend on interaction with which part of the event is highlighted by grammatical aspect. For instance, Ferretti et al. (2007) revealed that the imperfective grammatical aspect activates event-specific knowledge to prime common locations of events, in contrast to the perfect grammatical aspect. Location and instrument are more salient in ongoing events than in events completed. In English, the language tested by Ferretti et al., the imperfective is compatible with the ongoing part of an event; in their study, attention to the ongoing event part activated the event-specific knowledge. The English imperfective be V-ing has a progressive (e.g., John is running) or near-to-the-endpoint reading (e.g., John is arriving), but in either case the part of the event highlighted by the imperfective is some point before the end-point. In Japanese, however, the imperfective -tei-can have a progressive reading (with all durational verbs, i.e., non-achievement verbs) or a resultative reading (with achievement verbs [Shirai, 2000]), with the former highlighting the ongoing part of an event just as in English, while, crucially, the latter highlights the end-state. For instance, John-ga (eki-ni) tsui-tei-ta with the imperfective-tei- means that John had arrived (at the station), but not that John was arriving (at the station). Previous studies have intensively worked on the imperfective aspect when it has the progressive reading. But we still lack empirical data on the effect of the imperfective when it focuses on an end-state, which makes it incompatible with the ongoing part of an event. One possibility is that the attested effects of grammatical aspect may be confirmed regardless of which event part is in focus. Another possibility is that such effects may diminish when the imperfective aspect is not compatible with the ongoing part of an event. In this
research, I address the question of whether the effects of imperfective aspect reported in previous studies are to be ascribed to its interaction with event focus.

Furthermore, it has not yet been fully investigated whether the role of grammatical aspect is universal or language specific. Taking English and Japanese for examples, languages can be different in their word order, the semantics of their grammatical and lexical aspect (as discussed above), their structural complexity (e.g., Japanese allows scrambled word order, while English does not), and how children first produce grammatical aspect with certain verbs (the aspect hypothesis). In English sentences, a verb that is marked by tense and grammatical aspect appears earlier than it does in Japanese, where the verb marked by tense and aspect appears in the sentence final position. Structurally, in English the imperfective has the form of a *be* verb with tense and *V-ing*, while in Japanese it is more complex, technically consisting of three parts, *-te-i-ru*, where *-te* is a connective, *-i-* is a form of verb, and *-ru* marks tense. These differences may affect how we proceed in our sentence processing when comprehending situations described in language. Japanese-speaking children seem to initially mix verbs and tense-aspect morphology in unpatterned ways, while English-speaking children follow a pattern. These observations raise the question of whether the roles of grammatical aspect and the ways it interacts with lexical aspect are uniform across languages.

In this dissertation, I have outlined my research aims in Chapter 1. Chapters 2 and 3 discuss the existing theoretical and empirical research on language comprehension. Chapter 2 discusses the background of aspect in the literature on syntax and language acquisition. Chapter 3 focuses on previous psycholinguistic research that has detected
effects of grammatical aspect in sentence processing. In addition, in Chapter 3, I frame a set of questions on the effects on sentence processing of the interaction of grammatical aspect with lexical aspect and event duration.

Chapter 4 presents Experiment 1, which investigated a case where grammatical aspect is *incompatible* with lexical aspect. The experiment looked into readers’ attention to different event parts after they read a sentence in which lexical aspect and grammatical aspect were manipulated (e.g., *Evan was kicking/kicked a ball/balls to Megan. He/She_________*). Participants’ sentence continuations were collected in order to record their choice of referent, when one was relevant to the ongoing part and the other to the end state of the event. In this experiment, I am interested in whether grammatical aspect highlights a part of an event (i.e., the end state), given a “mismatching” combination of *atelic* lexical aspect and *perfective* grammatical aspect. In other words, the question this experiment addresses is whether grammatical aspect exerts its influence regardless of lexical aspect or depends on interaction with lexical aspect.

Chapter 5 presents Experiments 2 through 5, which reevaluate the cost of processing the imperfective, addressing the claim that cost depends on event duration (Coll-Florit & Gennari, 2011). I investigate whether such cost is solely due to grammatical aspect or event duration, or to their interaction. The prediction is, first, that imperfective sentences will induce longer reading times than perfective sentences, as Coll-Florit and Gennari found. But previous studies have suggested that the imperfective allows comprehenders more access to ongoing events and related world knowledge, so the prediction goes further, expecting that comprehenders’ reading times will reflect differences of event duration described as *short* or *long* when they read imperfective
sentences, but not when they read perfective sentences. I employed self-paced reading tasks controlling event duration and imperfective versus perfective aspect (Tom was golfing/golfed 18/three holes with his colleagues last Sunday) in English (Experiment 2) and in Japanese (Experiments 3 and 4). This chapter employed Japanese to address the question of whether the event part in focus influences the effects of grammatical aspect. Namely, Experiment 3’s materials used verbs that evoke the progressive meaning with the imperfective (as in English), while Experiment 4’s materials used other verbs, which evoke resultative readings of the imperfective. If focusing on the ongoing event part, as opposed to the end-state, is not crucial, then Experiments 3 and 4 would both find effects of grammatical aspect, but a diminishing effect would be observed in Experiment 4, where the imperfective focuses on the end-state. In addition, the word order difference in the two languages allows an interesting comparison: In Japanese, because grammatical aspect appears sentence finally, duration phrases come before grammatical aspect, but the opposite is true in English. Crucially, the processing cost of the imperfective aspect observed in Coll-Florit and Gennari’s study may be ascribed to what I would call openness. In other words, uncertainty about how long an event would continue may have made readers take longer to read imperfective sentences than perfective sentences. If the cost seen in the imperfective sentences is actually the cost of comprehenders’ wondering how long the event would continue, the expected longer reading times in the imperfective sentences may disappear if comprehenders obtain information on the event’s duration before they read the grammatical aspect. To test this, Experiment 5 removed the duration phrases from the Japanese sentences used in Experiments 3 and 4, so that comprehenders would not obtain the information on the events’ duration.
Previous research has also demonstrated that grammatical aspect can affect the ease of retrieving content from memory. Magliano and Schleich (2000) demonstrated that the imperfective makes a sentence’s information more accessible for memory retrieval compared with the perfective aspect. Their experiment used a sentence marked by the perfective or imperfective aspect, where a main character of a short story performed an action. The probe recognition latencies were collected at a close or far position during successive sentences, and the results revealed that the imperfective sentences induced faster reaction times than the perfective sentences, especially at the far probe position. In Chapter 6, Experiments 6 through 9 assess recognition latencies of a given verb (e.g., sleep) at close and far probe positions, after participants read sentences such as Atsushi was sleeping/slept in his room yesterday. These experiments test whether the grammatical aspect interacts with lexical aspect by manipulating verbs that evoke a progressive meaning with the imperfective (with the ongoing focus), and other verbs that evoke resultative readings of the imperfective (with the end-state focus). By employing Japanese as the test language, it is possible to explore whether grammatical aspect manifests an effect regardless of verb type and whether the way in which grammatical aspect highlights parts of an event affects memory retrieval of a described event. The prediction is that the imperfective will show its effects similarly with both the progressive and resultative readings if the imperfective, as a linguistic cue, plays a role uniformly across verb types, while, on the other hand, only the progressive will help to maintain an event representation in memory, as opposed to the resultative, if the highlight on event part is crucial to the role of grammatical aspect.
Chapter 7 concludes this dissertation by synthesizing the findings on grammatical aspect’s interaction with lexical aspect and event duration in English and Japanese. This reevaluation of the role of grammatical aspect in the framework of situation model construction will suggest that the effect of grammatical aspect is stronger than that of interaction, and that the effect of grammatical aspect on sentence processing may be susceptible to the timing of when aspectual cues are available to readers, leading to the discrepancy between Japanese and English reported in studies of children’s acquisition of these languages.

Few empirical studies have explored the interaction of lexical and grammatical aspect, and, to my knowledge, there have been no studies on processing outcomes from mismatching lexical and grammatical aspects, and no research on the role of grammatical aspect that highlights two separate parts of events with the same grammatical aspect marker (i.e., the imperfective -teiri-). This research contributes to our understanding of how we comprehend situations in language by introducing cross-aspectual and cross-linguistic investigation to the research on grammatical aspect.
CHAPTER 2

GRAMMATICAL ASPECT AND LEXICAL ASPECT
IN HUMAN LANGUAGE

2.1 Time and event in language

Human language encodes temporal properties of events by means of lexical and grammatical aspect. A verb (phrase) denotes situation types by its lexical aspect, also known as *aktionsart*, which encodes internal temporal properties of events in terms of durativity, telicity, and dynamicity. Based on these event properties that define aspectual classes, verbs have been classified into four verb classes: *state*, *activity*, *achievement*, and *accomplishment* (Vendler, 1957). Grammatical aspect, or “viewpoint” aspect, on the other hand, imposes a viewpoint onto an event; imperfective takes an inside viewpoint to see an event as on-going, while perfective takes an outside viewpoint to see an event as completed. Across languages, the major contrast in grammatical aspect is imperfective versus perfective (Comrie, 1976). For example, the progressive morphology is used to denote imperfective aspect and simple past morphology is used to denote perfective aspect in English.

2.2 Lexical aspect in syntax and semantics

Lexical aspect has been intensively studied by lexical semanticists interested in the temporal properties of events described by verbs or a verb and its complement(s). In earlier research, verbs were thought to form natural classes based on lexical aspect, which
was seen as inherent in lexical items.\(^1\) Verb classification based on temporal properties, or on temporal schemata (see Vendler, 1957), is motivated by the shared temporal aspectual properties of verbs or the events they describe, characterized as features such as \([\pm \text{durative}/\pm \text{punctual}], [\pm \text{dynamic}], \text{and } [\pm \text{telic}]\) (Table 2.1). The aspectual property of *duration* refers to whether a given event (inherently) goes on over a period of time (e.g., *walk*, *draw a circle* [+durative]), or whether it is punctual (e.g., *die*, *break the glass* [−durative]). *Dynamicity* distinguishes eventive verbs from stative verbs, implying movement or change, and that the event is not static. *Telicity* (called *terminativity* by Verkuyl [2000]) refers to an obligatory end-point that a given telic event has (e.g., *draw a circle, die*), in contrast to atelicity (called *durativity* by Verkuyl [2000]), which does not necessarily encode an end-point (e.g., *walk*, *sleep*). This property-based, or feature-based, approach classifies verbs into four semantic classes (e.g., in Vendler’s 1957 verb classification, often called Vendler/Dowty or Aristotle/Vendler classification), consisting of states (e.g., *know*, *love*), activities (or *processes*; e.g., *walk*, *sleep*), achievements (e.g., *die*, *explode*, *reach the top*), and accomplishments (e.g., *eat an apple*, *go to the store*). Smith (1991, 1997) recognized a fifth class, *semelfactives*, for punctual and atelic verbs (e.g., *jump*, *knock*, *bang*) that can refer to repetitive action that does not have a clear endpoint. Some scholars consider telic verbs (i.e., accomplishments and achievements) to form one class as *events* or *transitions* (Pustejovsky, 1991), which denotes a certain change or opposition. Pustejovsky (1991) claimed that any verb in natural language can

\(^1\) This idea is compatible with the lexical decompositional approach: The meaning of a word is claimed to be essentially decomposable into a set of features (Katz, 1972; Katz & Fodor, 1963; Schank, 1975; Wilks, 1975), or primitives (Dowty, 1979; Verkuyl, 2000) with mathematical principles (Jackendoff, 1972; Katz, 1972; Lakoff, 1970, 1971; Schank, 1975; Wilks, 1975).
be characterized as one of three basic event types: states, processes, or transitions (see also Dowty, 1979; Vendler, 1967). The classification presented in Table 2.1 is based on Vendler’s (1957) classification system.

<table>
<thead>
<tr>
<th>Verb classification</th>
<th>Dynamic</th>
<th>Duration</th>
<th>Endpoint</th>
</tr>
</thead>
<tbody>
<tr>
<td>state</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>activity</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>accomplishment</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>achievement</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>semelfactive</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
</tbody>
</table>

Lexical aspect, in the earlier history of its research, was originally thought of as a property of a lexical item, a verb, rather than a verb phrase (Dowty, 1979): Vendler’s quadripartition was thought to be lexically based. One can ascribe an event’s aspectuality to the lexical item of the verb. For example Jane ate a sandwich (telic) and Jane disliked a sandwich (atelic) form an aspectual minimal pair with only the verb being different. However, more recent semanticists have moved toward discussing lexical aspect as more compositional, rather than specified in a lexical item as a class of verb.

For instance, Verkuyl’s theory emphasizes that both the verb and its arguments contribute to the aspectuality of sentences (e.g., Verkuyl, 1972, 2000). For example, the contrast between the telic Jane ate a sandwich and the atelic Nobody ate a sandwich cannot be ascribed to the verb, which is identical in the two sentences, while the subject noun phrase (NP) makes them a minimal pair. Similarly, Jane ate a sandwich is telic while Jane ate sandwiches is atelic; again, the verb stays identical but the direct object expresses a bounded quantity on one hand (for a telic sentence) and an unbounded quantity on the other hand (for an atelic sentence). Verkuyl (2000, p. 2) addressed the
question of what basic semantic material underlies aspectual composition and how composition proceeds at higher phrasal levels. Verkuyl took the position of assuming that the meaning of a sentence is composed of the meanings of its parts. That is, interpretation is achieved by going from the bottom up, starting with lexical items and moving onto phrasal meanings. Verkuyl (2000) explicitly argued against Vendlerian (1957) classification because Vendlerian classification does not pay attention to the contribution of argument NPs. Aspectual interpretation can even depend on construction; that is, in some cases, aspectual meaning cannot be tracked down to the meaning of each word but to a construction. This is called aspectual coercion. For example, compare *The light flashed until dawn* with *The light flashed*. In the former sentence, the event of the lights flashing is durative and not punctual while in the latter, it is punctual. Aspectual coercion is considered semantically rich in the literature.

To summarize, more recent semanticists have discussed aspect’s compositional nature in relation to other constituents in a sentence. Such exploration has led semanticists to notice the strong association between aspect on the one hand and syntactic distribution and argument realization on the other; that is, the mapping between semantics and syntax in terms of aspect or event structure. Grammatical aspect has been less intensively discussed than lexical aspect in the literature, and often has been discussed additively.

### 2.3 Grammatical aspect in relation to lexical aspect

Lexical aspect, as we have seen above, has been classified in terms of aspectual semantic temporal properties as well as in relation to syntactic argument. Grammatical aspect is
less discussed than lexical aspect, and the discussion does not enter much into grammatical aspect’s role in constructing rich, internally structured lexical representations of aspect, although its role is to provide a viewpoint on a given event. For example, with the activity verb walk in *John was walking*, the imperfective aspect, expressed by *-ing*, provides a viewpoint internal to the event. The placement of such an internal viewpoint, however, is rather different with the achievement verb *die*, as in *John was dying*, where the viewpoint is on the temporal point right before the state of being dead. Croft (1998b) called the interpretation derived from achievement verbs in combination with imperfective grammatical aspect, which is not canonical progressive, *runup achievement*. Even with canonical progressives, as in *John was eating an apple* and *John was walking*, there are differences. In the former sentence, with an accomplishment verb, the internal viewpoint can depict various stages of *eat an apple*, while the later describes a *homogenous* event with an activity verb, where the depicted moments are homogeneous during the event.

Verkuyl (2000) does not treat grammatical aspect such as progressive as part of an aspectual theory but instead as part of a Tense system, in that it deals with events in connection with real time, such as whether the state holds while an event is in progress (and at no other time). The tense maker and grammatical aspect progressive are represented simply as functions: [INFL] and [PROG], respectively. Tense is a [INFL] function to take a tenseless S\(^2\) to make a tensed S’: S’ = INFL (S). The progressive form PROG applies before INFL and is syntactically defined as taking a tenseless S to form a tensed S’: S’ = PROG (S). On the other hand, Croft’s (2009) approach follows the

\(^2\) S = sentence
intuition that grammatical aspect imposes a viewpoint onto an event. An important element of Croft’s analysis is that it represents grammatical aspectual meaning by *profiling*. That is, it places the focus on which part of an event is *profiled* in the event representation. This approach considers causality and temporal structure as being of separate dimensions but interconnected, and views causal and aspectual structure as distinct levels (Croft follows the tradition of Dowty, and McCawley, 1971, by acknowledging causal chains, i.e., where a participant acts on another participant). At the level of aspectual structure, or a temporal/aspectual contour of each sub-event, represented are time (t) on the x-axis and qualitative states (q) on the y-axis.\(^3\)

(1) Source: Croft, 2009, p. 12

Each sub-event linearly represents the aspectual contour in the form of a line/dot/arrow that stretches horizontally from left to right, and these lines are vertically connected by lines/arrows to others in any other causal dimension(s) so that one can see at which temporal point a participant acts on another, making transmission of force visible. This *profiling* provides a simple representation for the unusual meaning in the runup achievement, such as *He’s dying*, where the interpretation is not canonical progressive, \(^3\)

If there is no causality, as in (1), the dimension for causation is not expressed.
but refers to the temporal point right before the state of being dead. In a representation, the bold parts in an arrow/line indicate the profiled part. For example, in the representation in (1) above, the black part of the line references the interval immediately before the state of being dead. A wavy line represents an undirected activity, where there is “not a measurable gradual change to the resulting state” (Croft, 2009, p. 12).

Such analyses as Croft’s, however, may only be descriptive, without explaining why the combination of lexical aspect and grammatical aspect generates such interpretations. The importance of the present discussion becomes more obvious when one considers how grammatical aspect is acquired by children and language learners across languages, where language learners are strongly aware of situation types as expressed by lexical aspect. The next section provides an overview of research findings on lexical and grammatical aspect in the language acquisition literature.

### 2.4 Aspect in child language

Early child language is interesting in that children, unlike adults, often show unique behavior in their language use. In the studies of language acquisition, it has been widely attested that young children produce tense-aspect morphology with specific verb types in a nonadultlike distribution that some researchers have interpreted to mean that children’s acquisition of tense-aspect morphology is guided by verbs. In this section, I will first discuss theoretical development around this topic. I will then introduce some research on Japanese-speaking children and second language learners in a following section. Japanese shows a different initial distribution between verbs and morphology than that observed across many other languages, including English.
Researchers in the field of language acquisition have long noticed a pattern frequently observed in early child speech. In a difference from adult grammar, children undergeneralize tense-aspect morphology, mapping specific morphemes to semantically well-defined subclasses of verbs. English-speaking children initially use past/perfective morphology\(^4\) predominantly with telic verbs (i.e., achievement and accomplishment verbs), and imperfective morphology predominantly with atelic verbs (i.e., action verbs) under the age of approximately 2;6 (Delidaki & Varlokosta 2003; Wagner, 2001). Such a consistent distribution of verb morphology in relation to specific verb types has been cross-linguistically attested in early child language production: English (Bloom, Lifter, & Hafitz, 1980\(^5\); Clark, 1996; Harner, 1981; McShane & Whittaker, 1988; Olsen & Weinberg, 1999; Shirai & Andersen, 1995), Italian (Antinucci & Miller, 1976), French (Bronckart & Sinclair, 1973), Polish (Bloom & Harner, 1989; Weist, Wysocka, Witkowska-Stadnik, Buczowska, & Konieczna, 1984), Turkish (Aksu-Koç, 1988, 1998), Japanese (Rispoli, 1981), German (Behrens, 1993), Portuguese (De Lemos, 1981), Mandarin (Li, 1990), Greek (Stephany, 1981, 1997), Cantonese (Erbaugh, 1978), and Hebrew (Berman, 1983). Akhtar and Tomasello (1997) explained that the patterns are “past tense inflections with verbs that describe nondurative kinds of events with clear end results, and the progressive aspect inflection with verbs naming durative events with no clear end state” (p. 962). Bickerton (1981, 1984) detected the state-process distinction and the punctual-nonpunctual distinction emerging in creole grammars (i.e., second

---

\(^4\) There is an issue of conflation in English, where the -\textit{ed} suffix marks both past tense and perfective aspect.

\(^5\) Bloom et al. (1980) described telic predicates as marked with past tense (rather than perfective aspect), and atelic predicates as marked with present tense (rather than imperfective aspect).
generation) even when the earlier pidgin (i.e., first generation) does not mark tense-aspect distinctions.

Along with the asymmetry observed in the distribution between verbs and tense-aspect morphology, there has been another asymmetry found in initial child language in the order of acquisition of the morphology (e.g., Akhtar & Tomasello, 1997; Brown, 1973; also see Behrend, 1990). Brown’s (1973) study showed that mastery of the English -ed morpheme occurs relatively late in development. In naturalistic data, the present progressive inflection -ing was used in 90% of obligatory contexts well before -ed, showing earlier mastery of the former than the latter. Akhtar and Tomasello (1997) provided experimental evidence from an elicited production task that children start productively using -ing before -ed: their participant group of two-year olds (2;1) added -ing to novel verb stems but did not show use of -ed, while the three-year-old group used both. Akhtar and Tomasello’s account of the late acquisition of -ed discussed, in addition to the role of phonological regularity and the relative frequencies of verbs that take these morphemes, the relevance of aspect versus tense to a verb’s meaning. In other words, importantly, Akhtar and Tomasello considered grammatical aspect to more or less directly affect the meaning of a verb stem but tense to apply to an entire proposition or clause (Bybee, 1985; Slobin, 1985a). This view is in line with that of O’Grady (2006), who pointed out that morphologically expressed grammatical aspect is marked closer to the verb root (Foley & Van Valin, 1984, p. 210), indicating that tense is dependent upon prior determination of aspect. Moreover, in languages where grammatical aspect is morphologically expressed, grammatical aspect is acquired either before or at the same time as tense, but never after tense (see O’Grady 2006; Van Valin 1991, p. 16).
Considering the localist hypothesis (Li & Bowerman, 1998; Tomasello, 1992) that morphology is interpreted on a verb basis, it is possible that aspect is more salient to children than tense.

Children’s use of aspect-tense morphology reflects a strong interaction between lexical and grammatical aspect. Grammatical aspect is even restricted in some languages to certain classes of verbs, such as perfective markers to telic and imperfective to atelic predicates in Mandarin (Li & Bowerman, 1998; Smith 1991). There is evidence that the interaction between lexical and grammatical aspect is innate. Olsen and Weinberg (Olsen 1997; Olsen & Weinberg 1999) accounted for the distribution in terms of the Subset Principle, which hypothesizes that children’s initial production starts with the most restricted relationship between grammatical and lexical aspect, constraining their use of -ed to telic verbs (i.e., [+telic]) and their use of -ing to dynamic and durative verbs (i.e., [+dynamic] and [+durative]). Olsen and Weinberg did not see evidence that children’s distribution was influenced by input from adults, and therefore argued that the distribution is based on Universal Grammar and that the role of adult input is only to provide positive evidence that eventually relaxes the initial restrictions. The idea is that children’s initial hypothesis is corrected in the development of language. Bickerton’s (1981, 1984) Language Bioprogram Hypothesis claims that certain aspectual distinctions are preprogrammed and innate, based on observations from pidgins and creoles. Bickerton argued that children use perfective/past tense morphology to mark “punctual vs. non-punctual [+/ punctiliar]” and use imperfective past or progressive morphology to

---

6 On the other hand, there are no languages that restrict tense marking according to lexical aspect.
mark non-punctuality, along with the distinction between “state vs. process,” for which children do not overgeneralize the -ing progressive marker with stative predicates.

Shirai and Andersen’s Aspect Hypothesis (Andersen & Shirai, 1996; Shirai, 1991; Shirai & Andersen, 1995; also see Bardovi-Harlig 1994a, 1999 for SLA) differs from Olsen and Weinberg’s Subset Principle explanation in that Shirai and Andersen suggested a prototype theory that incorporates the role of motherese and its co-relation with children’s tense/aspect production. Using Vendler’s (1957) verb categorization, they generalized the universal pattern of child tense-aspect morphology in relation to lexical aspect, as summarized below in (2). The distribution is considered to reflect the prototypes with which children start using certain inflections before extending their use to peripheral types (see also Li & Shirai 2000).

(2) 1. Children first use past (or perfective) marking predominantly with achievement and accomplishment verbs, eventually extending their use to activity and finally to stative verbs.

2. In languages that have progressive aspect, children first use progressive marking mostly with activity verbs, then extend it to accomplishment and achievement verbs.

3. Children do not incorrectly overextend progressive markings to stative verbs.

These observations support my intuition that there may be processing reasons for these patterns. For example, children may employ some processing mechanism that leads to the asymmetry observed in the distribution between verbs and tense-aspect morphology, or
adult speakers (i.e., motherese) may follow a process that eventually is reflected in children’s speech, or both child and adult speakers may be influenced by such a processing reason.

Researchers have assumed that children use inflection for some underlying reason. Seemingly opposing the idea of processing reasons (which may apply for both child and adult speakers), some researchers consider that the asymmetrical distribution in young children’s verb morphology reflects children’s confusion between aspect and tense: children may use aspect to learn tense. These ideas come from early proponents of the Primacy of Aspect (POA) hypothesis (Antinucci & Miller, 1976; Bloom et al., 1980; Bronckart & Sinclair, 1973), also referred to as the Aspect Before Tense Hypothesis (Bloom et al., 1980), Defective Tense Hypothesis (Weist et al., 1984), Redundant Marking Hypothesis (Shirai, 1993), or the Aspect First Hypothesis (Wagner 1998, 2001). The Aspect First Hypothesis also refers to the close association between tense-aspect morphology and lexical aspect itself (see Delidaki & Varlokosta, 2003). Shirai and Andersen (1995) capture the issue of distribution in the Aspect Hypothesis’s idea that children initially use lexical aspect/verbal morphology asymmetrically with verb types as prototypes and later extend their use of morphology to peripheral verb types (see also Andersen, 1991; Li & Shirai, 2000; Robinson, 1990). As for what children mark with the morphology, Shirai and Andersen (1995) describe the difficulty of knowing whether early past morphology encodes aspect or tense, as the prototypes of past (tense) and perfective (grammatical aspect) are very similar (Shirai & Andersen, 1995, p. 759). In my opinion, one concern may be the conflation of tense and grammatical aspect morphology in English (see footnote 4).
Other researchers emphasize cognitive reasons. As mentioned, the POA in its early version argued that children initially map lexical aspect onto tense. Early proponents of this idea included Bronckart and Sinclair (1973), Antinucci and Miller (1976), and Bloom et al. (1980). While some scholars suggested that children do this to learn tense, others ascribed it to the difficulty children have in distinguishing between present and past due to cognitive deficiencies (Antinucci & Miller, 1976; Bronckart & Sinclair, 1973). Thus, the distribution in child grammar shows children’s inability to distinguish present and past tense, which leads them to mark “more salient” aspectual distinctions instead of tense. Hyams’s (2007) claim that past, present, and modal meanings result from general principles of aspectual interpretation, in particular the Punctuality Constraint (Giorgi & Pianesi, 1997) seems to conform to the POA. Other possibilities of what children use the morphology for include mapping lexical aspect onto grammatical aspect (Olsen, 1997; Olsen & Weinberg, 1999) and mapping grammatical aspect onto tense (Wagner, 2001). The current postulation of the POA, presumably including the Aspect First Hypothesis (Wagner, 2001), is summarized by Slabakova (2002) thus: lexical aspect is mapped onto tense and grammatical aspect. The Aspect First Hypothesis explains that children may map lexical aspect, or grammatical aspect, onto tense morphology. For example, children may understand a verb with -ed to indicate “this is a telic predicate” and a verb with -ing to indicate “this predicate is atelic,” making a mapping between lexical aspect and the morphology. Alternatively, children may conceive -ed and -ing suffixes as encoding grammatical aspect: whether a given event is perfective (completed) or imperfective (ongoing).
Comprehension studies provide meaningful ways to test the Aspect First Hypothesis in terms of whether children map aspect (whether it is lexical or grammatical aspect) on tense. Recent studies have revealed evidence to support the claim that young children correctly comprehend tense (Valian, 2006), challenging the Aspect First Hypothesis if it proposes that children encode lexical aspect (not grammatical aspect) onto tense morphology. Wagner (2001) indicated that younger children (from 2;2 to 2;11, mean 2;6) use tense marking to encode grammatical aspect. For the two-year-olds, the meaning difference between *is* and *was* is linked to completion information; that is, they appear to use the auxiliaries to mark grammatical aspect. Thus, the Aspect First Hypothesis is viable if it considers grammatical aspect to be mixed with tense morphology and if it considers processing behavior for grammatical aspect to interact with lexical aspect.\(^7\) As we will cover in Section 3.3.1, recent findings in psycholinguistic studies show that acquisition and processing patterns fit together: past/perfective morphology with telic verbs (i.e., achievement verbs), and imperfective morphology with atelic verbs (i.e., action verbs) facilitate adult language processing.

Indeed, the evidence implies that the interaction between lexical and grammatical aspect plays a role in adult language acquisition. Although it once seemed that the association between verb type and tense-aspect morphology might reflect children’s

---

\(^7\) Some mixed results have been observed in experiments testing whether younger children interpret auxiliaries as conveying tense in an adult-like way. Delidaki and Varlokonta (2003) tested Greek-speaking children using the same methodology used in Wagner’s Experiment 2, with atelic-telic verb manipulation. Older children (from 3;2 to 4;0) showed understanding of tense better than younger children (from 2;7 to 3;1). Present tense posed no problem, but past and future tense seemed to be problematic particularly for the younger group. There were no significant differences in the use of telicity with past forms for both age groups, showing that lexical aspect did not play a specific role in tense morphology.
confusion about tense, the same distribution has been reported with second language learners. By and large, learners make initial associations between a prototypical semantic class of verb and the inflection, and then extend the inflection to less prototypical verbs (Collins, 2002, 2004; Duff & Li, 2002; Giacalone-Ramat, 2002; Housten, 1995, 2002; Rocca 2002). Specifically, the extension of the past/perfective marker from telic to stative verbs has been widely attested in a number of descriptive studies using different methodologies in the field of L2 studies (see Bardovi-Harlig, 1999, 2000; Collins, 2004; Salaberry & Shirai, 2002). Second language learners as adult speakers should know the concept of tense, and therefore their use of tense-aspect morphology cannot be showing a deficiency in knowledge of tense.

Collins (2004) claimed that it is a language learning universal that lexical aspect influences the acquisition of tense-aspect markers. In her study, Collins tested the influence, if any, of L1 (French and Japanese) on learning English as a second language. Unlike previous studies on L1 influence that looked at either a single L1 background (Collins, 2002) or bidirectional influence between Italian and English (Rocca 2002), this study chose participants whose L1s contrast in two specific ways. First, French has a form equivalent to the English perfect (the passé composé), while Japanese does not. Second, French does not have a form equivalent to the English progressive -ing, while Japanese has a functionally similar marker -tei-. Using a 25-passage rational cloze task, in which 56 out of 82 items targeted the simple past (the perfective aspect) for elicitation across the four verb types (i.e., states, activities, achievements, accomplishments), Collins confirmed that both French- and Japanese-speaking learners were most successful in the appropriate use of past tense morphology with telic verbs (achievements,
accomplishments), and least successful with stative verbs. Interaction between verb types and L1 was also observed in that the French-speaking learners were less successful (because they inappropriately used the present perfect) than the Japanese-speaking learners in the use of the simple past with achievement verbs. That is, L1 influence was observed only with achievement predicates. There were no differences between French- and Japanese-speaking learners in the patterns of their use of present forms with statives, or in their use of progressive forms with activities. Hence, the study found L1 influence to be limited and not sufficiently strong to counter the Aspect Hypothesis.

The question is why there is an asymmetrical distribution between lexical aspect and tense/aspect markers in early child and second language grammars. The distribution of verb-morphology may reflect innate distribution. Another possibility is a processing reason, that is, computational ease (O’Grady, 2006). O’Grady proposed that the computation of aspect creates difficulty for tense marking. O’Grady considered two types of memory system, declarative memory and procedural memory (e.g., Cohen, 1984; Cohen & Squire, 1980; Pinker & Ullman, 2002; Ullman, 2001). The former underlies knowledge and information associated with the lexicon, and the latter underlies grammar. O’Grady assumed that tense marking is an output of procedural memory. In his account, O’Grady referred to Gavruseva (2002a, b), who made a distinction between achievement verbs (for which aspectual status, i.e., of telicity, is inherently determined) and accomplishment and activity verbs (for which aspectual status is computed by reference to the type of complement that the verb takes). O’Grady assumed that tense marking

---

8 With activity verbs, although, overall, uses were no different between the two language groups, Collins noticed that the progressive was used more by lower-level Japanese learners, while Japanese speakers “somewhat surprisingly” produced perfect forms more often.
should be easier with achievement than with the other verb types that require additional syntactic computation. The potential evidence O’Grady cited includes the results of studies done by Andersen and Shirai (1996, p. 546–47), Rohde (1996), Shirai (2003, p. 204), and Collins (2004, p. 263).

Achievement verbs are well attested to be strongly preferred for past tense morphology (Andersen & Shirai, 1996; Collins, 2004; Rohde, 1996; Shirai, 2003). When discussing processing reasons, O’Grady considered ease of processing in terms of syntactic computation. Yet the semantic information of lexical aspect seems to pose interesting processing difficulties to investigate. Achievement verbs form a tricky verb class as they induce unique interpretations in either English or Japanese when marked by the imperfective aspect. In English, the imperfective with achievement verbs presents runup achievement (e.g., John is dying), while in Japanese it produces a resultative reading. John-ga (eki-ni) tsui-tei-ta with the imperfective -tei- means that John had arrived (at the station), but not that John was arriving (at the station), the progressive reading, as we will cover in Section 3.3.2. My research interest is in whether the possible underlying reason for children’s asymmetry of verb and tense morphology is any reason that involves semantics, which works on event representation.

2.5 Aspect in child Japanese

Shirai’s (1998) report indicated that the distribution between lexical aspect and tense/aspect morphology is not strictly consistent cross-linguistically. Shirai (1998) discussed data from three Japanese-speaking children that did not show the expected distribution of verb type and tense-aspect morphology. This is contrary to Shirai’s (1993)
earlier report of a strong association between the past form -ta and achievement verbs in data from a Japanese boy, Sumihare (Noji, 1976, 1977). In the later study, Shirai (1998) analyzed data from three Japanese-speaking boys, Aki, aged 1;5 to 3;0 (Miyata, 1995, from the CHILDES language database; MacWhinney, 1995; MacWhinney & Snow, 1985), Yocchan, aged 1;11 to 2;2 (Clancy, 1985), and Tacchan, aged 1;0 to 2;2 (Kokuritsu Kokugo Kenkyujo, 1982a, 1982b) for three types of marking: -ta (past) and two non-past markings, -tei- (durative) and -ru. Shirai saw a very early use of the past tense -ta with stative verbs with these children\(^9\). Note here that stative verbs are the least prototypical verb for the past/perfective acquisition according to the Aspect Hypothesis, as telic verbs are the prototype. Shirai reported that it was only after -ta became productive\(^10\) that the dominant use with achievement verbs was observed in the children’s data.

Shirai suggested potential accounts for the deviation from the predictions of the Aspect Hypothesis observed in these Japanese-speaking children’s initial verb-morphology distribution. Shirai mentioned the possibility that the children used -ta as a perfect marker, basing this suggestion on the historical fact that -ta come from the perfective -tari and is on its way to becoming a full-fledged past tense marker (Horie, 1997; Kinsui, 1997: Takahashi, 1976). Shirai reported that one child, Aki, appeared to use itta as a perfect marker, but could not use it in the obligatory past context though he used the form very early, indicating incorrect comprehension. Shirai also raised the possibility of a role being played by input from mothers’ speech. Shirai (1998) analyzed

---

\(^9\) Most of the instances reported by Shirai are the verbs of existence a-ru (for inanimate entities) and i-ru (for animate entities).

\(^10\) Shirai (1998) considered the use “productive” when the form in question was used with at least five different verb types during a particular month.
Aki’s mother’s speech, which showed a correlation between achievement verbs and the past/perfective -\textipa{\textit{ta}} (50.0–71.4\%, except for the period when the child was between 1;7 and 1;11). Aki showed the dominant use of achievement verbs with -\textipa{\textit{ta}} once it became productive. To summarize, Shirai’s study showed the possibility that child language could use -\textipa{\textit{ta}} to mark grammatical aspect, not tense, and that the distribution between verbs and the morphology might be influenced by the mother’s input. In a second language acquisition context, the perfect form was also reported by Collins (2004) to be used with unexpected frequency with a variety of verb types in the L2 English of Japanese speakers. In an English cloze task, the Japanese speakers showed this usage much more than French speakers, even though there is no counterpart of the perfect form in Japanese. The data showed much more use of past perfect than present perfect. Collins explained that the participants reported difficulty in understanding the difference between past tense and past perfect, and that this was an instructional effect: the Japanese participants were college or university students, and had most recently learned the past perfect, which was taught in their final year of high school.

The Aspect First Hypothesis again seems to be challenged by reports from Japanese, if the idea of the hypothesis is that tense is learned in accordance with a certain lexical aspectual property: Stative verbs do not represent events that are either punctual or completed. The idea that children encode punctuality or completedness with a tense maker is not viable with Japanese-speaking children. Child Japanese is not consistent in the use of verbs with tense-aspect morphology, unlike other child languages including child English. Therefore, if the reason that underlies children’s behavior in verb selection with tense-aspect morphology is one that is specific to children (such as to mark tense), it
will not be observed in adult Japanese. On the other hand, adult Japanese may show a similar pattern as child Japanese, and adult and child English may also share a pattern, if the reason that underlies children’s behavior in verb selection with tense-aspect morphology is a processing reason shared by children and adults. In the L2 English data provided by Collins (2004), two different first language groups (French and Japanese) uniformly showed effects of lexical aspect. Hence, it is possible that the sentence processing that underlies the verb and aspect use relies more on cognitive processing, and is related to event properties and representation, rather than being influenced by language differences (such as word order). To give an example of cognitive considerations, which will be discussed further in Section 3.3.1, Yap et al. (2009) suggested the similarity between lexical aspect (telicity) and grammatical aspect (boundedness). Having said this, the results of Collins’s study do not rule out a non-cognitive processing explanation, because the evidence came from an off-line study (i.e., a cloze task), which does not provide information about processing behaviors. In the next chapter, we will look into prior studies that focus on sentence processing in sentence comprehension.
CHAPTER 3

GRAMMATICAL ASPECT AND LEXICAL ASPECT IN SENTENCE PROCESSING: ROLES AND INTERACTION

Event representation, unlike mere construction of a representation of a text, is a dynamic enterprise performed in sentence comprehension. This idea is representatively captured as situation model construction, and researchers in psycholinguistics have detected that forms at various linguistic levels provide cues that contribute to event representation. Among such linguistic cues, the effect of grammatical aspect has been intensively explored in the framework of situation model construction. Yet it has been underdiscussed in relation to other sentential elements, such as the information about the event provided by lexical aspect. This chapter discusses findings from psycholinguistic studies that suggest what roles aspect can play in sentence comprehension in the framework of situation model construction. In addition, in this chapter I propose that by considering the interaction between grammatical and lexical aspect, as well as other sentential elements, researchers will be able to better interpret human sentence comprehension.

3.1 Situation model construction

Evidence has supported that comprehension of language involves construction of the state of affairs, or situations, that are described by a text. Such mental representations of linguistically described situations are known as mental models (Johnson-Laird, 1983) or situation models (van Dijk & Kintsch, 1983), and a research stream has developed that
focuses on sentence comprehension to address how comprehenders construct coherent situation models online, as construction of a coherent situation model is crucial for the successful comprehension of a text (e.g., Glenberg, Kruley, & Langston, 1994; Graesser, Millis, & Zwaan, 1997; Graesser, Singer, & Trabasso, 1994; Johnson-Laird, 1983, 1989; van Dijk & Kintsch, 1983; Zwaan, Magliano, & Graesser, 1995). The latest theory of situation model construction is the event-indexing model (e.g., Zwaan, Langston, & Graesser, 1995; Zwaan, Magliano, & Graesser, 1995; Zwaan & Radvansky, 1998; see also Speer, Zacks, & Reynolds, 2007; Speer & Zacks, 2005), which assumes five dimensions—time, space, causation, motivation, and protagonist—along which each event can be indexed. Comprehenders are thought to construct current situation models, updating them to reflect incoming story events of the situation (Magliano, Zwaan, & Graesser, 1998; Zwaan, Langston, & Graesser, 1995; Zwaan, Magliano, & Graesser, 1995; Zwaan & Radvansky, 1998). The time dimension is considered the most important dimension that comprehenders use in constructing situation models, and researchers have examined the roles of various levels of linguistic cues that concern time in event representation. These cues include grammatical aspect.

3.1.1 Situation model and grammatical aspect

Across languages, the major contrast in grammatical aspect is between imperfective and perfective (Comrie, 1976). The imperfective aspect imposes an inside “viewpoint” onto an event to see the event as ongoing, while the perfective aspect imposes an outside “viewpoint” to see the event as completed. Psycholinguistic researchers have examined whether this contrast plays a role in our online sentence processing. Studies with English
as the target language have employed the progressive -ing as the imperfective (in present or past tense), while for the perfective, some researchers have used simple past tense and others have used present or past perfect (have/had V-ed). This section discusses five important psycholinguistic findings on grammatical aspect and event representation.

One of the findings reported by psycholinguists is that grammatical aspect affects comprehenders’ event representation in that the imperfective aspect makes an event representation more accessible by having comprehenders interpret the event as ongoing and as part of the current situation model. For example, Carreiras et al. (1997, experiment 3) demonstrated that, in a short narrative with a target character (e.g., John) and another character (e.g., Mary), the target character was more accessible (i.e., subsequent probes using the character’s name produced faster responses) when the action was described in the past progressive (e.g., John was finishing his shift) than when it was described in the past perfect (e.g., John had finished his shift). Magliano and Schleich (2000) similarly showed that the effect of the imperfective grammatical aspect (e.g., was changing) is to make an event be considered more ongoing and thus more accessible for memory retrieval in the successive sentences, compared to the effect of the perfective aspect (e.g., changed). In their word recognition task, participants read a short story sentence by sentence, including a critical sentence that had a verb phrase marked either by the imperfective or the perfective. At some point after they had read the critical sentence, they saw a presentation of the verb phrase used in the critical sentence in nontensed form, and responded by pressing computer keys that were marked yes or no to answer whether they had previously read the word. The participants’ decision latencies were faster in the imperfective than the perfective. In addition, the difference of effects between the
grammatical aspects was larger when the probe was placed three sentences after the critical sentence than when it was placed immediately after it, indicating that the imperfective makes a sentence’s information more resilient for memory retrieval compared to the perfective aspect.

The second major finding on grammatical aspect reported in the psycholinguistic literature is that grammatical aspect affects the extent to which comprehenders’ event representation has access to world knowledge relevant to the described event. This finding strongly supports the argument that comprehension involves more than the construction of a representation of a text. Bergen and Wheeler (2010) showed that progressive sentences facilitated an action-sentence compatibility effect (ACE). They found that a progressive sentence about hand motion facilitated a manual action in the same direction as described in the sentence (e.g., *John is closing the drawer*), while a perfect sentence (e.g., *John has closed the drawer*) did not facilitate such an action. Their experiment examined the ACE by asking participants to read a sentence and to decide whether it was meaningful or not. Participants responded by pressing computer keys labeled *yes* or *no*. The keys were placed in such a way that answering required the participants to move their hand towards or away from their body.

A study by Zwaan and Stanfield (unpublished; cited in Zwaan, Madden, & Stanfield, 2001) showed preliminary results that support the role of grammatical aspect in activating world knowledge associated with a described action, in the form of both *semantic activation* and *episodic activation*. For instance, participants were given a story that introduced an instrument such as a hammer, but the subsequent sentences implied it was unavailable. Participant’ responses to *hammer* were found to be significantly slower
after they read a sentence such as *John began pounding the boards together* than after they read a sentence such as *John pounded the boards together*. Zwaan and Stanfield accounted for the results by positing that the imperfective descriptions were more likely than the perfective descriptions to activate the situation model that participants had constructed in which the instrument was not available, leading to *episodic activation* (in which an association is made between words mentioned and constructed in situation models). Moreover, even when the instrument was not mentioned at all, the recognition latency was slower after reading *began pounding* sentences than after reading *pounded* sentences. Zwaan and Stanfield interpreted this as *semantic activation* between *pounding* and *hammer*. Based on this observation, Zwaan et al. (2001) hypothesized that imperfective descriptions make more information about an event available than perfective descriptions. Although Zwaan and Stanfield’s manipulation of grammatical aspect employed *began V-ing* versus *V-ed* constructions, which may not represent the regular contrast of the imperfective versus the perfective, the experiment does provide evidence that grammatical aspect affects how comprehenders use world knowledge about described events in comprehending a sentence.

The third important psycholinguistic finding shows that the role of grammatical aspect in placing a “viewpoint” on a given event is psychologically real. Crucially, evidence has shown interaction between this role of grammatical aspect and other factors in sentence processing. For example, activated world knowledge, which was just discussed, may relate to a specific event part that is highlighted by grammatical aspect, and not necessarily to the entire event. Based on Moen and Steedman’s (1988) work, Ferretti et al. (2007) considered that entities, objects, and locations that participate in
events will be more or less salient at different junctures of the temporal and causal structure of events. For example, agents are associated with the initial part of an event, as they cause the event; patients are related more to the resulting states as they undergo the result caused by the agent’s action; and location and instrument are more salient during ongoing events than they are after the event is completed. Location may be background information compared to other participants, and in lexical decision tasks, Ferretti, McRae, and Hatherell (2001) did not see a priming effect between verbs and locations (skated–arena), which they did see between verbs and instruments (dug–spade), agents (arresting–cop), and patients(arresting–criminal)\(^{11}\). In a subsequent study, Ferretti et al. (2007) revised the test materials that included location, using four conditions that contrasted common versus uncommon locations and the perfective versus the imperfective (e.g., was skating–arena, was praying–arena, had skated–arena, and had prayed–arena). Ferretti et al. (2007) found that imperfective but not perfective grammatical aspect activates event-specific knowledge to prime common location of events (location is more salient in ongoing events). In the priming task, using a short stimulus onset asynchrony (SOA; the time between the onsets of the prime and the target), participants named the target location faster in the imperfective conditions. In addition, in their offline sentence completion task (experiment 2), participants completed the sentence fragments with prepositional phrases (PPs) more often after imperfective sentences (e.g., The diver was snorkeling _____ ) than after perfective sentences (e.g., The diver had snorkeled _____), and the proportion of locative PPs was higher in the imperfective

\(^{11}\) According to Ferretti et al. (2001), the past-tense form instead of the -ing form was used for the noun target because “an ing ending often resulted in an adjective-like pairing with the noun target (e.g., served–platter was used rather than serving–platter)” (p. 525).
condition than the perfective condition. The perfective condition induced more noun phrases for the sentence completions than the imperfective, which is consistent with the prediction that entities and objects are more associated with the result state. Ferretti et al.’s (2007) study hence shows evidence that grammatical aspect plays a role in modulating activation of world knowledge associated with a given event, particularly of a specific event part that is highlighted by the grammatical (i.e., viewpoint) aspect.

Ferretti et al. (2009) further presented evidence of the role of grammatical aspect in constraining readers’ attention to a specific entity that participates in a given event, by using verbs of transfer. Ferretti et al. (2009) tested sentences like John SOURCE handed/was handing a book to MaryGOAL, where thematic roles of source and goal participated in the events. They found that participants overall continued their sentences with the goal (e.g., Mary), showing a goal bias, but the bias was reduced in the imperfective condition more than the perfective condition (See also Rohde, Kehler, & Elman, 2006; and Ueno & Kehler, 2010, mirroring Ferretti et al. with Japanese participants). ERPs (event-related brain potentials) indicated that the readers experienced difficulty in processing a passage where the pronoun references the source noun rather than the goal noun with both grammatical aspects (e.g., John SOURCE handed/was handing a timecard to MaryGOAL. He/She asked about the upcoming meeting), but indicated that in the imperfective condition the cost takes place later than in the perfective condition.

Recent research has revealed not only the role of the imperfective aspect, but also the role of the perfective aspect, in the construction of situation models: The perfective aspect constrains comprehenders’ situation model to the end-state. It is worth noting that the previous evidence has been dominantly related to the imperfective aspect’s effects on
situation models. As for the perfective aspect, its role includes deactivation (e.g., Bergen & Wheeler, 2010), and it might be difficult to detect any effect of something that gets deactivated. However, by using a picture recognition task, Madden and Zwaan (2003) found an effect of the perfective aspect. Participants read the sentence John {made/was making} a fire followed by a picture that depicted the completion or ongoingness of the event; they then decided if the picture matched the given sentence or not. In the perfective conditions, participants’ correct responses were faster to completed pictures than to ongoing pictures, while no effect of progressive sentences was seen. This absence of imperfective effect is contrary to what we would expect, given the findings of the literature we have discussed. Madden and Zwaan suggested that it may have been because they used only accomplishment verbs, which describe nonhomogenous, culminating events. Culminating events can have various ongoing stages, and their images may vary from participant to participant. For example, the ongoing part of building a house (an accomplishment verb) can refer to the stages where materials are collected, a house is half built, a house is almost built, and so on, while images for completion would vary less, referring only to the point of completion of the house. Therefore, the pictures used to depict ongoing events might have been ambiguous, and might not have matched the participants’ mental images. If this reasoning is correct, the imperfective aspect with activity verbs, which express homogenous events (e.g., walk, sleep), would elicit an effect in such a picture recognition task. Such an effect was confirmed by Yap et al. (2009) in Cantonese. In a similar picture task as that used by Madden and Zwaan (2003), but introducing a factor of verb type (accomplishment vs. activity), Yap et al. confirmed that both grammatical aspect types constrained
comprehenders’ situation models to particular event components: the imperfective to ongoing parts, and the perfective to event end-states. Crucially, the studies that came before Yap et al. had not explicitly controlled verb types for the grammatical aspect, nor grammatical aspect for the lexical aspect, and had individually tested the effect of each type of aspect.

The fourth role that grammatical aspect plays in sentence processing is related to the *iconicity assumption*. The iconicity assumption has been defined to mean that a comprehender’s default assumption is that the narrated order of events is expected to match their chronological order (e.g., Hopper, 1979). In the *strong* version, the comprehender assumes no time gaps between the narrated events. The empirical evidence shows that comprehenders experience a processing cost when the iconicity assumption is overridden by linguistic cues such as *before/after, then, in an hour*, which indicate a time leap, and grammatical aspect. For instance, Munte, Schiltz, and Kutas (1998) tested the effect of chronological or reversed expressions (i.e., ‘before’ and ‘after’ sentences) by recording ERPs (event-related brain potentials) and showed greater negativity for *before* as early as 300 milliseconds into the comprehension of this sentence-initial word. Bestgen and Vonk (1995) found that temporal markers indicating explicit versus implicit cues of a time gap between the described events, such as *and/then/Ø*(no adverbial), can affect the availability of situation models to memory. That is, the *and* and the absence of a temporal marker made previous information more available than a sequential marker, *then*, supporting the *strong* iconicity assumption. Time shifts implied by cues such as *{A moment/An hour/A day} later* (Zwaan, 1996) have been shown to affect online comprehension. However, this was only so after explicit cues of discontinuation of events
such as stopped walking, as opposed to was walking (Zwaan et al., 2000). In other words, research results show that time shift can lead to slow event retrieval, supporting the strong version of the iconicity assumption, and time shift can be overridden if the event is considered ongoing, as cued by the imperfective grammatical aspect.

The fifth finding on grammatical aspect in sentence processing is that imperfective sentences are reported to be more costly than perfective sentences. Coll-Florit and Gennari (2011) showed that, in Spanish, readers take longer to read imperfective than perfective sentences. Their phrase-by-phrase self-paced reading task compared reading times for sentences such as *El portero tapaba el cartel* ‘The door man was covering the sign’ and *El portero tapó el cartel* ‘The door man covered the sign’, and found longer reading times for sentences of the former type at both the critical region (verb with grammatical aspect, *tapaba*) and the post-verbal position (*el cartel*). The observed cost was discussed as showing event “duration,” considering the imperfective sentence to describe a durative event while the perfective sentence describes a punctual event. This idea was supported by their participants’ perceptions of the length of time of the events described by the test sentences. Participants rated the duration (on a scale from 1 to 7) of, for example, the situation described by *El portero tapaba el cartel* ‘The door man was covering the sign’ as lasting longer than the situation described by sentences such as *El portero tapó el cartel* ‘The door man covered the sign’.

To summarize, psycholinguistic studies have revealed that grammatical aspect influences language comprehension by addressing how grammatical aspect affects comprehender’s construction of situation models. The imperfective aspect keeps an event representation accessible by making it be interpreted as ongoing as part of a current
situation model. The imperfective aspect highlights different particular event components than the perfective aspect. Some recent studies have shown that both the imperfective and the perfective aspect seem to play a role in constraining comprehenders’ attention to particular event components. Imperfective sentences are costlier to process than perfective sentences. However, the core element that conveys event information should be the verb (and verb phrase) and its lexical aspect.

3.2 Psycholinguistic studies on lexical aspect

The core semantics for event interpretation is encoded in verb phrases, which is to say in lexical aspect, not in grammatical aspect. Though it has been explored by a relatively smaller number of studies, research has started to show that lexical aspect also has a certain influence on online language processing. This includes effects from telicity (Husband, Stockall, & Beretta, n.d.; Malaia, Wilbur, & Weber-Fox, 2009; O’Bryan, 2003), duration (Coll-Florit & Gennari, 2011), and dynamicity (Gennari & Poeppel, 2003). The psycholinguistic research on lexical aspect has detected the role of lexical aspect in syntactic processing, or in the cost of processing due to the complexity of semantic structure, but lacks studies on situation model construction.

For example, O’Bryan (2003) revealed a hint of the role of verb telicity and transitivity in helping a comprehender to process a syntactically misleading sentence (i.e., so-called garden-path sentences) created by a reduced relative clause (RRC), such as The actress {'awakened (telic) / worshipped (atelic)'} by the writer left in a hurry. With four conditions created by manipulating telic versus atelic and obligatory versus optionally transitive verbs as the verbs in the RRC, a self-paced reading task showed telic verbs
tending to reduce processing difficulty more than atelic verbs at the disambiguation point (e.g., by), and obligatorily transitive verbs tending to reduce the difficulty more than optionally transitive verbs at the main verb (e.g., left). O’Bryan’s study showed numerical differences between the conditions, but, because the conditions were unbalanced, the findings were not supported by statistical analyses such as ANOVA. Malaia et al. (2009) replicated O’Bryan’s results and further supported them by recording event-related brain potentials (ERPs), showing that telic verbs induce significantly less difficulty in syntactic reanalysis of RRCs than atelic verbs. Malaia et al. manipulated the same four conditions as O’Bryan’s, and the telic verbs all described events that contained a completed change. Malaia et al. ascribed the cost of atelic verbs to the necessity for thematic role reassignment: atelic verbs initially assign both agent and undergoer thematic roles to the first argument (e.g., The actress), but when readers notice that the verb is actually transitive and not intransitive, the readers need to reassign the roles between the subject and the object of the RRC (e.g., the worshipper and the worshippee). Telic verbs, on the other hand, can alternate between their noncausal (intransitive) and causal (transitive) interpretation while preserving the thematic role interpretation of the object (e.g., the actress is the one who is awakened): the additional argument is added as an external argument (e.g., the writer) without the need to reassign thematic roles, resulting in less processing cost. If we were to account for this observation in terms of situation model construction or event-index model construction, the processing costs might be ascribed to the required updates on the model; fewer updates would need to be made for awakened than for worshipped, as the former does not require an update on how to interpret the thematic role of the object, while the latter does.
Other studies on event semantics ascribe higher processing cost to the complexity of causal structures (Gennari & Poeppel, 2003; McKoon & Macfarland, 2000, 2002). For instance, Gennari and Poeppel (2003) compared semantic complexity in event structure between *stative* (noncausal) and *eventive* (causal) verbs. (In the Vendler-Dowty classification, stative verbs are state verbs and eventive verbs are activity, accomplishment, and achievement verbs.) According to Gennari and Poeppel, eventive verbs (e.g., *destroy, discover, play*) entail one or several changes and cover an additional CAUSE or BECOME function, hence are more complex than stative verbs (e.g., *love, belong, contain*), which entail a single stable situation without adding additional functions. The results confirmed the prediction, showing slower reading in a self-reading task and slower responses in a word/nonword recognition task for eventive verbs than for stative verbs. McKoon and Macfarland (2002, see also 2000) used a distinction between externally caused change of state verbs (e.g., *break* needs two participants) versus internally caused change of state verbs (e.g., *grow* needs one participant), rather than the Vendler/Dowty verb classification that is based on temporal aspectual properties. McKoon and Macfarland found that externally caused change of state verbs resulted in longer reaction times than internally caused change of state verbs in acceptability judgment, sentence reading, and word/nonword recognition tasks. They ascribed the costs of externally caused verbs to the complexity of their event structure, “α CAUSE [x [BECOME IN STATE]],” compared to the internally caused change of state verbs’ event structure, “x [BECOME IN STATE].”

In the studies conducted by O’Bryan (2003), Malaia et al. (2009), McKoon and Macfarland (2000, 2002), and Gennari and Poeppel (2003), the role of lexical aspect
seems to be explored in terms of the event structure, which in some cases refers to aspectual properties and in other cases more to causalities than to aspectual properties. To date, the role of lexical aspect in the construction of situation models has not been addressed. A possibility is that the costs found in the previous lexical aspectual studies may reflect costs of constructing situation models. For example, the higher costs of processing stative verbs than dynamic verbs, or external causative verbs than internal causatives, might be due to difficulties in constructing situation models. In other words, constructing situation models may require more updates with eventive verbs (because they entail changes) than with stative verbs (which entail no changes), or more updates with external causative verbs (because they entail more participants) than internal causative verbs (which entail no additional participants). Furthermore, the costs of processing lexical aspect may dynamically change in relation to sentential context, because grammatical aspect has been shown to highlight particular properties and parts of a given event, such as placing more focus on the ongoing part of an event in imperfective sentences than in perfective sentences, and comprehenders have been shown to have access to world knowledge regarding focused parts of events.

3.3 Exploring interaction between grammatical and lexical aspect

This review of psycholinguistic studies on aspect and sentence processing reveals gaps in the evidence for the role of temporal cues in sentence comprehension. Less attention has been paid to sentential elements in discussions of the effect of grammatical aspect. For instance, attested roles of grammatical aspect have been relatively less discussed in relation to what it modifies, i.e., event and its lexical aspect. The role of lexical aspect in
language processing is not directly related to the theory of situation model construction, which research on grammatical aspect has explored. With various aspectual temporal properties that characterize distinctive behaviors of verb classes, and/or with verbs characterized as entailing various event structures in terms of the (causal) relations of subevents inside an event, it is natural to presume that the information from lexical aspect would contribute in one form or another to language comprehenders’ construction of situation models. The discrepancy in the psycholinguistic literature is that research questions have been targeted at examining the role of grammatical aspect alone, or the role of lexical aspect alone, and studies are usually conducted on grammatical aspect and lexical aspect independently. Researchers’ interests are not much shared between these research streams, with situation model construction on one hand discussed in grammatical aspect research, and with processing cost and/or syntactic resolution on the other hand in lexical aspect research. In our daily language use, however, language comprehension is usually performed where lexical aspect and grammatical aspect together mark events. I therefore suggest that psycholinguistic studies have left open the important question of whether and how interaction effects exist between lexical and grammatical aspect.

3.3.1 Interaction of lexical and grammatical aspect in mismatching combinations

I take an interactional view on aspect effects, in that interaction between lexical and grammatical aspect exists, and the effect of grammatical aspect should not be discussed independently of lexical aspect. The processing literature has started to provide evidence that certain combinations of lexical and grammatical aspects work well for processing
(Yap et al. 2006, 2009). In Yap et al.’s picture verification tasks, participants were asked to listen to a sentence and respond to a picture that followed, deciding whether a given picture described a sentence. With four conditions of two verb types (accomplishment vs. activity) by two grammatical aspect types (perfective vs. imperfective), the results showed “perfective facilitation” in the accomplishment-perfective combination, and “imperfective facilitation” in the activity-imperfective combination. Yap et al. account for the observed facilitation by ascribing it to what they call like reinforces like (2009, p. 592). That is, similar features of lexical and grammatical aspect reinforce each other: the telicity feature that refers to an end-point in an accomplishment verb matches the boundedness feature in the perfective, while the durativity of an activity verb matches the ongoingness feature of the imperfective. However, Yap et al. overlooked the important question of how comprehenders deal with a “mismatching” combination: they failed to address what exactly takes place in a mismatching combination of aspects.

One theoretical speculation discussed by Madden and Ferretti (2009) is that such a mismatching combination of lexical and grammatical aspect would lead to aspectual coercion. Aspectual coercion is a function that the semantic system employs “when the grammatical aspect of a verb is at odds with that verb’s inherent temporal semantics” (Madden & Ferretti 2009, p. 14), as in the case of The light flashed until dawn as opposed to The light flashed. The idea I suggest here is that when the grammatical aspect has a certain interpretation (e.g., completion status from the perfective), and the situation expressed by lexical aspect does not match (e.g., is without an end-point), the grammatical aspect nonetheless may force a perfective interpretation onto the situation.
Aspectual coercion assumes that a powerful role is held by grammatical aspect. One might say that this idea of aspectual coercion conforms to the previous empirical studies on grammatical aspect’s effect, where researchers tested its effect without controlling verb types and considered the effect of grammatical aspect to be the same with all verb types. However, aspectual coercion at the interface between grammatical and lexical aspect has not yet been empirically demonstrated. I consider it possible that coercion at the interface between grammatical and lexical aspect is implausible; I hypothesize that grammatical aspect will not lead to its expected interpretation if there is no certain information of aspectuality in a given lexical aspect. More specifically, I propose the following hypotheses.

**Hypothesis**: Endpoint is necessary in lexical aspect for the perfective to induce an interpretation of event completion.

**Hypothesis’**: Because endpoint is necessary for the perfective to induce a “completed” interpretation, no “completed” interpretation will be available when there is no endpoint.

These statements explicitly deny the existence of aspectual coercion at the interface between lexical and grammatical aspect. The above hypotheses lead to the following predictions.

**Prediction**: An unbounded event with the perfective does not lead to a “completed” interpretation (hence no coercion).

**Prediction’**: Because unbounded events with the perfective do not lead to an interpretation of completion (no coercion), unbounded events stay “unbounded” in processing. Thus, this may be similar to the case of the imperfective aspect, which provides an “internal view” of an event that it modifies, in which an end-point is out of view.
First, in Chapter 4, I test whether there is any coercion called for when there is a mismatch between lexical and grammatical aspect, as has been discussed in the literature.

### 3.3.2 Interaction of lexical and grammatical aspect in sentence processing, and understudied potential effect from focusing change

The interactional view on aspect effects should also be applied in the discussion of processing costs. Coll-Florit and Gennari (2011) demonstrated that imperfective sentences took readers longer to process than perfective sentences. However, it has not been explored whether the costs come from grammatical aspect alone or from interaction between grammatical and other temporal cues such as lexical aspect, which describes an event through the verb (phrase). Some evidence has implied that world knowledge about an event (e.g., event duration) contributes to comprehenders’ conception of the event duration. Magliano and Schleich (2000, Experiments 1 and 2) used a self-paced reading task with a short story containing a target sentence marked with either the imperfective or perfective, followed by a test question in various positions across three sentences after the target sentence. The questions were designed so that an answer of “yes” implied that the comprehender considered the activity as completed. Experiment 2 explored the effect of world knowledge of a given event, addressing whether knowledge about an event that is considered short (e.g., *doing 5 push-ups*) or long (e.g., *golfing 18 holes*) would affect the participants’ behavior in Experiment 1. Along with the main effect of grammatical aspect, with the imperfective making comprehenders consider the event more ongoing than the perfective, an interaction between aspect and world knowledge (i.e., event duration) was observed: Participants considered short-duration activities marked by the imperfective as
still ongoing at earlier sentence positions but not at later sentence positions (i.e., activities were no longer seen as ongoing later in the timeline of the story), while there was no difference between the sentence positions with long-duration activities. These results show grammatical aspect interacting with comprehenders’ world knowledge of events. Anderson, Matlock, and Spivey (2010) presented another piece of evidence that the effect of grammatical aspect interacts with a described event and comprehender’s world knowledge (e.g., spatial information). They claimed that differences in processing between imperfective and perfective aspect are differences in the prominence of certain semantic properties. To test this claim, they observed how comprehenders moved a character (by moving a computer mouse) on a terrain depicted on a computer screen, after hearing a sentence that described a terrain as hard or easy to navigate, such as The road to the university was {rocky and bumpy/level and clear}, followed by an aspect sentence like David {was walking/walked} to the university where he sat in class (past progressive vs. simple past). The results showed that the duration of the mouse track on a depicted path was longer in the context of hard terrain than of easy terrain when the participants heard imperfective sentences, while there were no differences in duration times when they heard perfective sentences. That is, the participants’ behavior reflected their use of world knowledge in comprehending a described event when they heard the sentences in the imperfective aspect but not when they heard the perfective sentences. These results accord with Magliano and Schleich’s (2000) conclusions. Previous research, therefore, has presented evidence that the imperfective aspect interacts with event knowledge in sentence processing, such that the imperfective makes event representation, as ongoing and internal, more accessible to comprehenders, while the perfective
represents events as completed, and only the end-state seems to be accessible to comprehenders.

Empirical studies have shown that comprehenders’ attention is driven to ongoing parts of events with imperfective sentences and to the end-state with perfective sentences. The highlighted part of an event has been reported to impact information activation (Ferretti et al., 2007, 2009); the world knowledge that is accessed is related to the prominent event part. However, the case has never been tested where the imperfective grammatical aspect highlights an event part other than ongoing, incomplete event parts.

In Japanese, the imperfective aspect can highlight not only an ongoing event part but also the end-state, as an outcome of interaction between lexical aspect and grammatical aspect. The Japanese durative grammatical aspect marker -tei-has drawn much attention from scholars for its ability to allow a resultative interpretation (3c) rather than the progressive readings (3a) and (3b), as the following examples illustrate.

(3)

Activity: action in progress
a. Ken-ga utat-te i-ru.
   Ken-NOM sing-ASP-NPST
   ‘Ken is singing.’

Accomplishment: action in progress
b. Ken-ha isu-o tukut-te i-ru.
   Ken-TOP chair-ACC make-ASP-NPST
   ‘Ken is making a chair.’

Achievement: resultative state
c. Asoko-ni booru-ga oti-te i-ru.
   there-LOC ball-NOM fall-ASP-NPST
   ‘The ball is there as a result of having fallen.’
Compare: iterative action-in-progress

d. Ken-wadoa-o tatai-te i-ru.
   Ken-TOP door-ACC bang-ASP-NPST
   ‘Ken is banging on the door.’

Shirai (2000, p. 332)

* List of abbreviations: ACC = Accusative, ASP = Aspect, LOC = Locative, NOM = Nominative, NPST = Nonpast, TOP = Topic

For instance, *John-ga (eki-ni) tsui-tei-ta* with the imperfective *-tei-* means that *John had arrived (at the station)*, but not that *John was arriving (at the station)*. It is known that the resultative reading is usually available with achievement verbs (those with an end-point but almost no duration, such as *arrive, open*),\(^{12}\) and with other verb types it induces progressive readings as in English sentences (Shirai & Kurono, 1998). The resultative use of *-tei-* is compatible with the end-state of an event, not with an ongoing part of an event, as opposed to its progressive use. The imperfective aspect can highlight either an ongoing part (just as the imperfective does in English), or the end-state (just as the perfective does in English and Japanese). This raises the question of whether the imperfective plays its role invariably, regardless of the focus on event parts. Previous studies relate grammatical aspect’s effect to what event component is highlighted (e.g., Ferretti et al., 2007; Madden & Zwaan, 2003); but a full exploration requires a study on grammatical aspect that includes these Japanese cases.

In Chapter 5, I investigate interaction between grammatical and lexical aspect in terms of two main issues. First, I intend to confirm whether comprehenders’ world

---

\(^{12}\) Although the achievement verbs provide one typical condition in which the resultative reading is available for the imperfective *-tei-*, Japanese linguists have claimed that defining the conditions where the resultative reading is available for the imperfective is complex (Kindaichi, 1950; Kudo, 1995; Okuda, 1978a, b; Shirai, 2000; see Shirai, 2000 for discussion).
knowledge of a given event (i.e., of event duration) is reflected in processing costs in the expected direction in the imperfective sentences but not in the perfective sentences.

Second, I will explore whether such expected interaction is uniform regardless of the prominence of different event parts in the imperfective sentences; that is prominence determined by a combination of lexical and grammatical aspect. Grammatical aspect may not exert its effect equally on an entire event in its representation; instead, it may possibly function in interaction with the specific part of an event that is highlighted by grammatical aspect. The exploration is on the role of grammatical aspect, especially imperfective aspect, asking whether the attested effect in the previous studies is an effect of the imperfective aspect alone or in interaction with other sentential elements.

3.3.3 Interaction of lexical and grammatical aspect in resiliency of event representation in memory

The interaction between lexical and grammatical aspect has to be confirmed generally with other attested effects of grammatical aspect in situation model constructions. In memory retrieval, Magliano and Schleich’s (2000) study showed that the imperfective event is more resilient in memory than the perfective counterpart. Other studies have shown that sentences evoke world knowledge about specific event parts in relation to the viewpoint given by the grammatical aspect (e.g., Ferretti et al., 2007, 2009). One synthesized assumption is that in event representation, certain focused parts of events would be particularly resilient in memory, considering that comprehenders have more access to highlighted event parts. If the attested effects of grammatical imperfective aspect are uniform (i.e., no interaction is found between lexical and grammatical aspect),
then in Japanese, where the imperfective aspect can highlight the end-state (typically with achievement verbs) in its resultative reading, it should be possible to confirm the resiliency of event representation with the imperfective aspect. An issue, however, is that Magliano and Schleich considered an event representation marked by the perfective aspect, which gives prominence to the end-state, to be less resilient because the focus is on the completion of the event.

In Chapter 6, I assume a possible factor that may change the effect of grammatical aspect on the construction of situation models: the interaction between lexical and grammatical aspect that induces differences in prominence in event representation. The chapter focuses on resiliency in memory to explore aspect’s effect on situation models, by conducting experiments equivalent to those conducted by Magliano and Schleich (2000), while introducing Japanese imperfective versus perfective sentences to test cases where the interaction is predicted to cause change by its unique combination of aspect and highlighted event parts.

In the next chapters, therefore, to address these questions on the interaction of grammatical aspect, I draw on a synthesis of evidence from prior research on grammatical aspect, lexical aspect, language acquisition, and adult sentence processing. Working through the experiments, I aim to assess whether the effect of grammatical aspect reported in the previous studies is an effect of grammatical aspect alone or can be ascribed to its interaction with other sources of temporal cues. By testing a unique form of interaction between lexical and grammatical aspect in Japanese, I further assess whether highlighting specific event parts contributes to the effect of grammatical aspect.
These experiments present one of the few investigations of the interaction of grammatical and lexical aspect.
CHAPTER 4

ROLE OF GRAMMATICAL ASPECT WITH MISMATCHED LEXICAL ASPECT IN CONSTRUCTION OF MENTAL EVENT REPRESENTATIONS

Studies have shown that grammatical aspect functions to constrain comprehenders’ event representations to a certain stage of the described event. Recent works have revealed that certain combinations of lexical and grammatical aspect, when similar properties of lexical and grammatical aspect are matched, can facilitate the construction of event representations (Yap et al., 2009). However, it is not known how, or even whether, grammatical aspect functions given a “mismatching” lexical aspect. Section 3.3.1 briefly discussed two hypotheses: In the first, readers construct mental representations of completed events whenever the perfective grammatical aspect is used to describe an event; in the second, readers will not construct such representations when perfective grammatical aspect appears along with verbs that lack matching lexical aspect, i.e., telicity. Let us call these the Coercion Hypothesis and the Non-Coercion Hypothesis, respectively. It is difficult to tell which hypothesis is supported by the results of studies that explore behaviors like reaction times, such as Yap et al.’s picture selection study, because aspectual coercion is reported to be costly (Piñango, Zurif, & Jackendoff, 1999; Piñango, Winnick, Ullah, & Zurif, 2006) and any latency in reaction times would be compatible with either hypothesis.

This chapter presents a sentence continuation experiment that was conducted to test how atelicity of lexical aspect affects comprehenders’ expectations of which
individual would be mentioned next with transfer-of-possession verbs. In contrast to the research done by Yap et al. (2009), dependent measures are explicit in the present study, which compares how often participants choose a Goal referent or a Source referent. Which referent is chosen is related to the readers’ interpretation of the event as completed or not completed (i.e., ongoing). The methods employed are taken from the experiments conducted by Ferretti et al. (2009), who tested sentences like John {carried/was carrying} a tray to Mary. They found that participants overall preferred to continue the sentence with the referent Mary, showing a bias towards the end-state of the transfer (Mary). This bias was reduced in the imperfective condition compared to the perfective condition, indicating that the end-state is less accessible to participants in the imperfective condition. It is of note that Ferretti et al.’s (2009) study used only telic verbs: transfer-of-possession verbs describe an event that has a natural termination when the theme entity has reached a goal. But in the present study, crucially, the distinction of telicity vs. atelicity was manipulated by the plurality of the direct object noun. I predicted that the bias towards the end-point of the transfer-of-possession event would be reduced if lexical aspect plays a role in the functioning of grammatical aspect, such that for a sentence like John carried trays to Mary, readers’ bias towards the end-point (Mary) would be reduced because the event’s end-state is unspecified. Alternatively, a bias towards the end-point in such a sentence would indicate aspectual coercion.

4.1 Perfective grammatical aspect and lexical aspect without endpoint

4.1.1 Participants
Eighty-one students at the University of Hawai‘i at Mānoa participated in an about 20-minute long experiment for the compensation of course credit or of a small bag of snacks. All were native speakers of English and some also spoke other language(s) natively. The dominance of English in the non-mono-lingual speakers was self-reported as 50% or more.

The data were compiled from two separate participant pools, which took part in the experiment at two different periods; I call the earlier experiment Experiment A (36 participants) and the later one, Experiment B (36 participants). These two experiments consist of Experiment 1 in this dissertation. There are some minor differences in the methodology of Experiments A and B, which will be described in the following sections.

4.1.2 Materials

Critical Stimuli. The stimuli consisted of 60 test items adapted from 72 test items used by Ferretti et al. (2009) in their Experiment 2. No fillers were included in this sentence completion study to reduce participants’ fatigue. Each test item was a sentence that included a transfer-of-possession verb with a Source and a Goal referent. The distinction of telicity vs. atelicity was manipulated by varying the plurality of the direct object noun. For example, *John handed books to Mary* has an interpretation that the event of handing a book takes place more than once (repetitively), without a specified end-point, making the event atelic. In contrast, *John handed a book to Mary* has an interpretation that the handing of a book was a definite single occurrence. Three lists were created based on a Latin-square design, and the stimuli in each list were presented in random order. Each participant saw stimuli in which one third of the verb phrases had perfective aspect with a singular noun as the direct object (4), one third had perfective with a plural direct object.
(5), and one third had imperfective with a singular direct object (6). See Appendix A for a complete list of critical sentences used in Experiment 1.

(4) Condition 1
Stimulus: Perfective-Telic: Evan kicked a ball to Megan.
Target: {Evan/Megan/He/She}__________.

(5) Condition 2
Stimulus: Perfective-Atelic: Evan kicked balls to Megan.
Target: {Evan/Megan/He/She}__________.

(6) Condition 3
Stimulus: Imperfective-Telic: Evan was kicking a ball to Megan.
Target: {Evan/Megan/He/She}__________.

The Source and Goal referents in the stimuli were both proper names and pronouns that differed in gender. The gender of the referents was balanced across stimuli. Also, the gender of the first-mentioned (i.e., Source) referent was balanced. The stimulus sentences were all identical between Experiments A and B. The Source or Goal referents for participants to continue the sentences with were either proper names or pronouns in Experiment A with proper names in 36 items and pronouns in 24 items. In Experiment B, the Source and Goal referents for all 60 items were expressed in pronouns. Proper names and pronouns were unambiguously bound to the Source or Goal referents by different genders, i.e., he or she. The modification of using all pronominalized referents in Experiment B was employed because post-experiment interviews with participants

13 Among the three lists for Experiment A, one list was later detected to have three items (out of 60) with a miscoding in the order of mention in the preamble sentence and the order of presentation in 2AFC for the two referents. Coding was fixed for analyses but slightly changed the distribution of items in List 1, from 30:30 to 27:33.

57
revealed that they had a tendency to continue their sentence with the familiar proper names regardless of the referent types. Therefore, Experiment B was designed to avoid this issue by using pronouns only.

Furthermore, care was taken in Experiment B to remove any effects of correlation between the order of mention of the referents in a given stimuli and the order of presentation of the referents in the two-alternative forced choice (2AFC; see Section 4.1.3 for a description of 2AFC). In Experiment A, items using pronouns and those using proper names were fixed items, i.e., they were of between-item design. The order of presentation of the two referents in the 2AFC was also fixed, and half of the items had the same order for the 2AFC as for the referents in the stimuli, while half had the reverse order. The implementation for Experiment B was modified such that the combination of the order of mention in the stimuli and the order of presentation in the 2AFC was always randomized in such a way that for half of the items, the order of mention in stimuli and the order of presentation in the 2AFC were the same while for the other half they were reversed.

**Iterative Likelihood Norming.** The items with plural nouns used in these experiments were intended to have a repetitive interpretation, as has been discussed in the lexical semantic literature (e.g., Verkuyl, 2000). However, there was a possibility for certain items that the plural referent could be interpreted as a collective theme, with a reading where the event of transfer takes place with all the collective referent at once, rather than with one referent at a time (i.e., repetitive). In such a case, the interpretation would no longer be atelic but telic. This possibility was evaluated by conducting a norming study.
Eight participants viewed the items in the plural-noun-with-perfective condition (5), and they were instructed to read sentences and select either the collective interpretation (i.e., “whether the event takes place all at once”) or the repetitive interpretation (i.e., “whether the event takes place one at a time”) based on their intuition. The repetitive likeliness was calculated for each of the 60 items based on how many participants considered the item repetitive, on a scale from 0 to 1, with 0 being robustly unlikely repetitive (i.e., robustly collective), and 1 being robustly repetitive.

The norming study confirmed that participants overall showed a mild preference for a repetitive interpretation. Out of the 60 items, 33 were scored .75 or higher and 15 items were scored .25 or lower, with the average being .64.

**Co-Located Likeliness Norming.** Rohde, Kehler, and Elman (2006) conducted Goal vs. Source experiments similar to those done by Ferretti et al. (2009), but also employed semantic properties of co-location of Goal vs. Source, so that participants had to consider whether Source and Goal were co-located. The present study also takes this factor into account in the analyses. For this purpose, co-locatedness was evaluated by conducting a norming study, where 11 participants viewed the items in the singular-noun-with-perfective condition (4), and were instructed to read sentences and select either co-located if they considered the Source and Goal referents to be in the same place, or not co-located if they considered the Source and Goal referents to be in different places, based on any criteria of their own. The co-locatedness was calculated for each of the 60 items based on how many participants considered the items as co-located, with the scale ranging from 0 to 1 with 0 being robustly not co-located and 1 robustly co-located.
The norming study confirmed that, overall, the participants (6 monolingual
speakers of English and 5 bilingual speakers of English) showed a mild preference for a
colocated interpretation. Out of all 60 items, 42 were scored at .73 or higher, while 18
items were scored at .64 or lower, with an average of .72.

4.1.3 Procedure

Participants were instructed to read a sentence presented by itself on a computer screen
and to type in their own continuation sentence in the text box. Before typing in the
sentence, participants had to perform a two-alternative forced choices task (2AFC) to
select the Source or Goal referent of the given sentence to continue their own sentence
with. The two referents, each in a pronominalized or a proper-name form, were presented
side by side on the screen, when participants hit the space key after having read the
stimulus sentence. Participants selected one referent by pressing keys labeled L or R.
Pressing a key for the 2AFC triggered the presentation of a text box, in which participants
typed in the rest of their sentence. For example, after being presented with a target
sentence *Evan kicked a ball to Megan*, participants saw *Evan* and *Megan* on the screen.
They would select one of these by hitting L or R, and then type in, for example, *was
unhappy* in the text box. Hitting the enter key terminated the sentence creation for that
item and moved to the next item. The instructions to participants were given visually and
orally: Participants were instructed orally while seeing written instructions on the screen.
The presentation of stimuli was in random order across items and participants, and
measurements were controlled and collected by E-prime software 1.2.
4.1.4 Conditions

Experiment 1 tested three conditions with two variables: grammatical aspect and telicity of verb phrase. The condition of interest has a mismatch between the grammatical aspect of perfective, and telicity based on the plurality of the noun, leading to one factor (telicity) with two levels (telic vs. atelic). These conditions were compared to the conditions used in previous studies, such as those of Ferretti et al. (2009), that manipulated grammatical aspect with a telic event, i.e., one factor (grammatical aspect) with two levels (perfective vs. imperfective). These factors are combined to make three conditions: telic perfective, atelic perfective, and telic imperfective.

4.1.5 Predictions

The present study was conducted to test the two particular hypotheses in which an atelic verb may, or may not, lead to an interpretation of event completion when mismatched with the perfective.

*Non-Coercion Hypothesis.* Since a compatible property for the perfective, the end-point, is not specified by lexical aspect in a mismatching combination of atelic with perfective, comprehenders’ attention should be less on the end-state and participants should interpret an event as not completed, hence this combination is expected to show less Goal bias than the matching combination of telic with perfective. If this is the case, it cannot be accounted for by Ferretti et al. (2009), who claim that grammatical aspect has the same effect in combination with all verb types, without considering verb types as contributing factors. This study replicates Ferretti et al.’s study but with the modification of
manipulating the plurality of nouns. The prediction is that Source references will be significantly more frequent following the imperfective aspect than following the perfective aspect. It is important to the present hypothesis that the baseline condition (telic with the imperfective) shows no difference in the proportion of Goal vs. Source bias than the condition of interest (atelic with the perfective). This baseline condition, with the imperfective, showed significantly less Goal bias than the telic with perfective condition in Ferretti et al.’s study. In the present study, the condition of interest (atelic with the perfective) should show similar results as this baseline condition, i.e., reduced Goal bias in comparison to the telic with perfective condition.

Coercion Hypothesis. According to this hypothesis, grammatical aspect functions no matter what lexical aspect it co-occurs with; therefore, comprehenders’ attention should be on the end-state even in a mismatching combination of atelic with the perfective, and hence the condition should not result in a reduction of Goal bias. Source references will be significantly more frequent following the imperfective than following the perfective, regardless of telicity.

4.1.6 Data analysis

Selected Participants and Items for Source Ratio. Although the present study was not designed to explore response times (RTs) in the selection of a Source or Goal referent, they were nonetheless examined to check for outliers as well as to confirm whether there are any relations in patterns between RTs and the Source or Goal preferences.
Data from nine participants were eliminated because of extreme latency of response times (RTs) in the 2AFC, with such latency being considered an indication of inattentiveness of participants; even if participants were attentive, such long RTs mean that their responses may not reflect participants’ initial reactions to the referent words. Outliers in RTs were trimmed at a fixed cutoff point of 5000 milliseconds (ms) for this first data cleaning, with RTs that were 5000 ms or more replaced with 5000 ms. Participants more than 20% of whose RTs were replaced by 5000 ms were removed. Based on this criterion, four participants from Experiment A and one participant from Experiment B were removed. In order to have an equal number of participants for each condition, an additional two participants in Experiment A and another two participants in Experiment B were removed. As a result, data from 72 people (i.e., 12 participants in Experiment A and 12 participants in Experiment B for each of three conditions) were left for further analysis.

Data Cleaning for Response Times. On the 72 participants’ data left for the source ratio analyses, further data cleaning was conducted for the decision time analyses. The original raw reading times were re-examined. Obvious outlying response times were first removed, so that response times that were over 8000 milliseconds were replaced with the cutoff value, with the idea that these times represented obvious outliers that would skew the subsequent mean value and standard deviation calculations. This resulted in 21

14 RTs were also examined by item. One item in Experiment A exceeded 20% (i.e., 21% of RTs from Item #35: to push a shopping cart), and was replaced with 5000ms. However, as the value was close to 20% and there were no other items that exceeded 20% in Experiment A or B, all 60 items were kept to maintain consistency and statistical power.
response times, i.e., 0.50% of RTs, being replaced on the entire set of trials. Second, outlying participants were sought by calculating the mean value and standard deviation of the replaced RTs of the entire set of trials for each participant, with the idea that participants should be excluded if their mean RT was below or above 3 standard deviations of the overall mean. Overall means were calculated separately for Experiments A and B, and outlying participants were examined with the applicable overall mean separately in Experiments A and B. One participant (mean: 3080) was removed from Experiment B due to the overall mean criteria (the average overall mean RT from 36 people in Experiment B was 1627 ms), but no participants were removed from Experiment A (the average overall mean RT from 36 people in Experiment A was 1900 ms). Third, outlying items were found by calculating the mean value and standard deviation of reading times in the items from all participants, with the idea that items should be excluded if their mean RT was above or below 3 standard deviations of the overall mean. Again, overall means were calculated separately for Experiments A and B, and outlying items were examined with the applicable overall mean separately in Experiments A and B. One item was removed (#35: to push a shopping cart) from Experiment B due to the overall mean criteria (the average overall mean RT in Experiment B from the 60 critical sentences was 1585 ms), but no items were removed from Experiment A (the average overall mean RT from the 60 items in Experiment A was 1900 ms). Finally, in order to have an equal number of participants for each condition, two more participants were removed from Experiment B based on the largest reaction times. As a result, 69 people’s data were left (i.e., 12 participants in Experiment A and 11 participants in Experiment B for each of three conditions) for further analysis.
**Selected Response Times.** The new mean value and standard deviation of the reaction times (with 8000 ms replacement) were taken for each participant. Trials were replaced if their RT was below or above 3 standard deviations with the cutoff value. With this criterion, 2.2% of overall raw RTs on the entire set of trials were replaced.

### 4.1.7 Results

**Source Ratio.** Mean source preference ratios\(^{15}\) computed across participants are plotted in Figure 4.1 for each condition. The columns in the graph show the combined results from Experiments A and B, and the bars show individual results from Experiments A and B.

As Figure 4.1 shows, Source bias was overall low across all the conditions, indicating that participants generally preferred the Goal referent for the first word in their continuation sentences. One-sample \(t\)-tests were used to compare the Source ratio to a hypothetical mean of 0.5 (i.e., 50%); Participants presented Goal bias following perfective context sentences with either a singular noun (by participants, \(M = 0.235, SD = 0.1764; t_1(71) = -12.76, p < .0001; t_2(59) = -17.57, p < .0001,\) two-tailed) or a plural noun (by participants, \(M = 0.2431, SD = 0.1753; t_1(71) = -12.41, p < .0001; t_2(59) = -18.22, p < .0001,\) two-tailed), as well as following imperfective context sentences (by participants, \(M = 0.2833, SD = 0.185; t_1(71) = -9.94, p < .0001; t_2(59) = -14.84, p < .0001,\) two-tailed). This result was similar to that of a similar production experiment by Ferretti et al. (2009, Experiment 1), where they reported Goal

\(^{15}\) Throughout the analyses, Source-preference ratio was calculated by dividing raw counts of Source referent by the sum of raw counts of Source referent and raw counts of Goal referent.
bias (85.3% following perfective context sentences and 77.3% following imperfective context sentences), which also appeared in a subset of sentences in which the reference to the Source or Goal was pronominalized (74.8% following perfective context sentences and 62.7% following imperfective context sentences) in their free continuation task. In contrast, in a study by Rohde, Kehler, and Elman (2006) that used an ambiguous pronoun prompt for the first word of the continuation, the Source bias was much higher, ranging from 51% to 70% in the continuation task.

However, the strength of the Source bias in the present study differed crucially by grammatical aspect (conforming to Ferretti et al., 2009; Rohde, Kehler, & Elman, 2006, among others), and not by interaction between lexical and grammatical aspect. Paired t-tests (alpha level at .05) on the results from Experiments A and B combined revealed that
Source references were significantly more common following the imperfective condition (with a singular noun), at 28.33%, than in either perfective condition; that is, either with the singular noun, 23.47%, both by participants and by items ($t_1(71) = 3.15, p = 0.0024$; $t_2(59) = 2.81, p = 0.0068$, two-tailed) or with the plural noun both by participants and by items, 24.31%, ($t_1(71) = -2.37, p = 0.020$; $t_2(59) = -2.62, p = 0.011$, two-tailed). Further statistical tests showed similar patterns even when Experiments A and B were separately examined, although Experiment B showed stronger effects while Experiment A showed weaker effects.

Contrary to the non-coercion hypothesis that predicts that interaction of lexical and grammatical aspect should affect the bias, the perfective condition with a plural noun did not change the bias. In the perfective conditions, it was hypothesized that Goal bias would be reduced (i.e., Source bias would be increased) when the perfective aspect is combined with an atelic event (e.g., *to kick balls*), which lacks an end-point, being a mismatch of aspectual property and perfective aspect. Paired $t$-tests on the combined results from Experiments A and B indicated that Source references did not differ between the mismatched perfective condition with the plural noun, 24.31%, and the matched perfective condition with the singular noun, 23.47%, either by participants or by items ($t_1(71) = 0.54, p = n.s.; t_2(59) = -0.51, p = n.s.$, two-tailed). Further statistical tests did not change the patterns, even when Experiments A and B were separately examined.

Further detailed sub-analyses took into consideration other factors that may have affected the results, such as iterativeness, co-locatedness, and use of pronouns vs. proper names for referents in 2AFC, but the overall pattern of results was not significantly affected by any of these factors. See Appendix B for a complete set of these sub-analyses.
**Response Times.** As mentioned, the present study was not designed to explore response times in the 2AFC, but they were nonetheless examined to explore whether there were any relations in patterns between RTs and Source preferences. Mean decision times in selecting Source or Goal referent computed across participants are plotted in Figure 4.2 for each condition: the times were plotted from the combined results of Experiments A and B, and individually from Experiment A (with 36 participants) and Experiment B (with 33 participants).

![Figure 4.2. Mean decision times by participants in selecting Source or Goal referent after reading a Perfective-Singular, Perfective-Plural, or Imperfective-Singular preamble, and 95% confidence intervals for each condition in Experiment 1](image)

<table>
<thead>
<tr>
<th>Condition</th>
<th>Perf-Sg(Telic)</th>
<th>Perf-Pl(Atelic)</th>
<th>Imp-Sg(Telic)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A+B</td>
<td>1669</td>
<td>1702</td>
<td>1737</td>
</tr>
<tr>
<td>Expt A</td>
<td>1805</td>
<td>1920</td>
<td>1928</td>
</tr>
<tr>
<td>Expt B</td>
<td>1521</td>
<td>1465</td>
<td>1528</td>
</tr>
</tbody>
</table>
Overall, the only meaningful difference in decision times across conditions was observed between imperfective-singular and perfective-singular in Experiment A. Paired t-tests were conducted on each pair of three conditions: perfective-singular vs. perfective-plural, perfective-plural vs. imperfective-singular, and imperfective-singular vs. perfective-singular, separately for Experiments A and B, and for the two experiments combined. The results suggested that participants made their decision on a referent (regardless of whether it was Source or Goal) more slowly after they read imperfective-singular sentences (by participant, $M = 1927.89$, $SD = 622.15$) than after they read perfective-singular sentences (by participant, $M = 1805.48$, $SD = 512.35$) ($t_1(35) = 2.18$, $p = 0.036$; $t_2(59) = 2.12$, $p = 0.038$, two-tailed). This effect appeared in the combined results from Experiments A and B as well, but it was weaker ($t_1(68) = -1.86$, $p = 0.068$; $t_2(59) = -1.89$, $p = 0.064$, two-tailed). Further analysis of the observed effect in Experiment A was conducted by dividing Source from Goal preference. The results suggested that, only in the by participant analysis, slower decisions in selecting a Goal referent were induced by the imperfective-singular (by participant, $M = 1879.91$, $SD = 631.42$) than by the perfective-singular (by participant, $M = 1751.74$, $SD = 559.87$) ($t_1(35) = 2.085$, $p = 0.045$; $t_2(59) = 1.39$, $p = 0.17$, two-tailed), but there was no difference between the two conditions when a Source referent was selected. There was another suggestion, only in Experiment A, that perfective-plural slowed down decision times (by participant, $M = 1920.20$, $SD = 564.85$) compared to the perfective-singular only by participants ($t_1(35) = -2.78$, $p = 0.0087$, two-tailed) and marginally by items ($t_2(59) = -1.83$, $p = 0.072$, two-tailed). No tests are reported for the rest of the comparisons, as the conditions did not differ in these pairs.
A Two-Factor ANOVA with repeated measures with a factor of aspect and with Experiments A and B as a between factor showed that RTs in Experiment A were significantly slower (by participants, $M = 1884.52$ ms, $SD = 541.24$) than in Experiment B (by participants, $M = 1504.37$ ms, $SD = 291.52$) both by participants ($F_1(1, 67) = 12.84$, $p = 0.00064$) and by items ($F_2(1, 67) = 103.67$, $p < .0001$). There was a suggestion of interaction between aspect and Experiment groups, but it was only significant by participants ($F_1(2, 134) = 3.17$, $p = 0.045$), and was marginal by items ($F_2(2, 234) = 103.67$, $p = .0080$). There was no main effect of aspect. Because Experiment A used both proper names and pronouns of Source and Goal referents in 2AFC while Experiment B used only pronominalized forms, the longer RTs observed in Experiment A may suggest a repeated names penalty (Gordon et al., 1993), which will be discussed in the next section.

4.1.8 Discussion

Experiment 1 was a sentence continuation experiment using the *verbs of transfer* to test how atelicity of lexical aspect affects comprehenders’ expectations of which individual would be mentioned next. Stevenson, Crawley, and Kleinman (1994) found that comprehenders showed Goal bias in verbs of transfer sentences (in the ambiguous pronoun prompt condition), despite expectations based on previous claims for Source bias in ambiguous pronoun resolution (subject preference by Crawley, Stevenson, & Kleinman, 1990; first-mention privilege by Gernsbacher & Hargreaves, 1988; grammatical parallelism preference by Smyth, 1994 and Chambers & Smyth, 1998). Succeeding studies, such as that done by Rohde, Kehler, and Elman (2006), further
separated Stevenson et al.’s explanation of thematic role prominence from the event structure explanation (also see Ferretti et al., 2009; Kim, Grüter, & Schafer, 2013, testing Korean; Ueno & Kehler, 2010, testing Japanese). Essentially, Stevenson et al. used only perfective aspect, but Rohde, Kehler, and Elman showed that Goal bias was decreased in the imperfective grammatical condition. They explained that the Goal referent is associated with the end-state, while the imperfective aspect mismatches the end-state, and thus the Goal bias was decreased in the imperfective condition. The present study predicted that the perfective condition would reduce Goal bias more with atelic verbs than with telic verbs, because an atelic event lacks an end-state. The present study was hence especially designed to test the hypothesis that an atelic verb would theoretically not lead to an interpretation of event completion even in combination with the perfective (non-coercion hypothesis). Namely, the prediction was that (5), like (6), would result in less Goal bias than (4), indicating that comprehenders’ attention would be less on the end-state, and showing that comprehender’s interpretation is that an event was not completed.

(4) Condition 1
Stimulus: Perfective-Telic: Evan\textsubscript{SOURCE} kicked a ball to Megan\textsubscript{GOAL}.
Target: \{Evan/Megan/He/She\}_________.

(5) Condition 2
Stimulus: Perfective- Atelic: Evan\textsubscript{SOURCE} kicked balls to Megan\textsubscript{GOAL}.
Target: \{Evan/Megan/He/She\}_________.

(6) Condition 3
Stimulus: Imperfective-Telic: Evan\textsubscript{SOURCE} was kicking a ball to Megan\textsubscript{GOAL}.
Target: \{Evan/Megan/He/She\}_________.
Contrary to this prediction, however, the results suggested that grammatical aspect is independent of lexical aspect information. This study’s results demonstrated a Goal bias with verbs of transfer across all the aspect conditions (conforming to the claims of Stevenson et al., 1994), but the Goal bias was affected by grammatical aspect, regardless of the use of atelic or telic events.

Several factors other than aspect may have influenced the observed Source bias, but in this discussion I focus on the possibility that the results show evidence of the influence of aspectual manipulation. First, the form of the reference to Source or Goal referents may influence the bias. For example, using unambiguous pronoun referents (as in the present study and Ferretti et al.’s 2009 study), instead of using ambiguous pronominal referents (such as in Stevenson et al.’s 1994 study), may have increased the Goal bias. Ferretti et al.’s study and the present study had high Goal bias (more than 50%) in both the imperfective and perfective conditions, and that was much higher than in Rohde, Kehler, and Elman’s (2006) study (less than 50%). Ferretti et al. explain that the ambiguous pronoun reduces this Goal bias, while in the unambiguous pronoun condition where participants can signal their referent, they showed similar Goal bias as in the proper name condition, in which Goal bias is reported to be strong (Arnold, 2001; Stevenson et al., 1994). Rohde and Kehler (2008) also confirm this, using a bare-prompt condition. A large Goal bias was observed with referring expressions, and this bias was diminished when an ambiguous pronoun prompt was provided in an ambiguous context. The high Goal bias in our study was similar to that in Ferretti et al.’s study. Possible mechanisms that might be at play are the subject preference (Crawley, Stevenson, & Kleinman, 1990), the first-mention privilege (Gernsbacher & Hargreaves, 1988), and the
grammatical parallelism preference (Smyth, 1994; Chambers & Smyth, 1998), which all prefer an ambiguous pronoun to coreference with the subject antecedent of the previous sentence, i.e., the Source. These mechanisms might explain a relative increase of the Goal bias in the unambiguous pronoun situation.

Second, using a proper name to refer to the entity in the subject position in the previous sentence may increase Goal bias more than using pronouns. Ferretti et al.’s results showed effects from the repeated name penalty (Gordon, Grosz, & Gilliom, 1993). That is, it was expected to be harder for comprehenders to refer to the entity in the subject position in the previous sentence by repeating names. By appealing to Centering theory (Grosz, Joshi, & Weinstein, 1995), Ferretti et al. (for their free-prompt experiment) hypothesized (i) that a large majority of the references to the (subject) Source should be pronominalized, (ii) a large majority of named references should be to the (non-subject) Goal, and (iii) both of these percentages should not differ significantly across the aspect manipulation (as this manipulation only influences next mention biases, and not grammatical role positioning or the relevant centering constructs). All these expectations were borne out in their study. Instead of the free prompt conditions used in Ferretti et al.’s study, the present study forced participants to choose between a pair of referents that were already both proper names or both pronouns. In such a situation, where participants have no choice in deciding to use either a proper noun or a pronoun, it is difficult to apply the same predictions. However, regarding (i), one can expect that participants may find it difficult to choose proper name referents in 2AFC when they were to select a Source
The results from decision times in 2AFC in Experiment A revealed that participants indeed took longer when referents were expressed in proper names, compared to all pronoun situations in Experiment B. It is worth reiterating, though, that the present study was not intended to collect time measurements. A study that was intended to collect decision times would be designed differently, having considered, for example, how to avoid any effects on decision times, such as the time that was spent by participants in the current study to think about the contents of the sentences that they were going to create. Still, such decision times on referent selection add behavioral evidence on how use of proper names or pronouns can influence the level of difficulty people have during coreferential processing, because those times were not collected in earlier offline story completion studies, although one study (Ferretti et al., 2009, Experiment 2) collected event-related brain potential (ERP) data to provide temporal information in a reading task for coreferential processing.

Third, another factor that could explain the referent bias in sentence continuation is the discourse coreference relation (Rohde, Kehler, & Elman, 2006). I believe, however, that it is unlikely that a discourse coreference relation could be responsible for the referent bias in the current data. An earlier study, by Kehler (2002), argues that discourse coherence relations underlie a variety of heuristic pronoun interpretation strategies that have been proposed in the psycholinguistic literature, and that they are side effects of the process of establishing coherence relations. This claim was illustrated by discourse examples, such as Miriam sent a fruitcake to Rachel. She didn’t want to eat it. If a reader

16 Note that this discussion is separate from the question of why the overall Source vs. Goal preference is not different depending on whether the 2AFC included proper names or not. Prior to the 2AFC, participants already had a preference, which is that they may find it difficult to use a proper name form when their preference is for the Source.
considers that the second clause explains the first, then the pronoun is most likely Miriam. If a reader takes the second clause as occasioned by the first, then the pronoun is most likely Rachel. An occasion relation is a relation where the first eventuality preconditions the second one. In other words, it requires that “the initial state of the second eventuality be identified with the final state of the first” (Rohde, Kehler, & Elman, 2006). In the present example, Miriam’s sending a cake to Rachel preconditions Rachel’s reaction of not wanting to eat it. Rohde et al. revealed that such occasion relations were dominated by continuations that refer to a Goal referent, while elaboration and explanation relations\textsuperscript{17} showed no evidence of Goal bias. Ueno and Kehler (2010, p. 2060) replicated this in Japanese, revealing that “Elaboration and Explanation are highly Source-biased whereas Occasion and Result are highly Goal-biased.”

One cannot deny a degree of difference, though, in the potential effects that a coherence relation may have on preferences regarding the referents. Rohde et al.’s (2006) study had two referents of the same gender in the first clause, which was followed by an ambiguous pronoun prompt that could be bound to either referent. Hence, in such a study, the reader may have played a greater role in the pronoun resolution decision than in this study. The present study, on the other hand, had unambiguous pronouns that referred to entities in the first clause. Ueno and Kehler (2010), like Rohde et al., used two referents of the same gender and controlled three prompt types, Null pronoun, Overt pronoun, and Free (no) prompt, where sub-analyses showed no difference between imperfective and

\textsuperscript{17} Elaboration and explanation relations are discourse relations in which the second clause adds more information onto the first clause.
perfective in Null and Free prompts.\textsuperscript{18} Therefore, for the most part, I consider that discourse coherence is rather an outcome than a cause behind the preferences for Source vs. Goal in this line of studies. Especially in the present study, I take the view that, in reading the first clause, the Source or Goal referent may become salient to readers due to factors such as the linguistic cue of perfective vs. imperfective; the reader constructs a story about the referent, and in this process discourse coherence is shaped by a language producer.

After all, Source vs. Goal preference explained by discourse coherence relations is another piece of evidence of aspect effect. For example, explanation/elaboration relations are discourse relations in which the second clause adds more information onto the first clause, and the occasion relation assumes the completion of the event described in the first clause. It is natural that the occasion relation is established when the first clause is the perfective. And Goal bias is explained by the end-state’s salience with the completion of the event, rather than by discourse coherence. That is to say, discourse coherence itself does not explain why Goal bias is increased in occasion relations, as the occasion relation can co-exist with a Source referent, too.

Having said this, if discourse coherence plays more or less of a role in such referent bias, the perfective-plural condition may be no different in its effects than the perfective-singular condition. A limitation of the present study is that the atelicity is partly telic; linguistically, the whole event \textit{to kick balls} is atelic, but part of this event can be carried out such that a ball or two may have been kicked to Mary, indicating that the end-state is partly available, which allows the occasion relation, even when, linguistically,

\textsuperscript{18} More specifically, Free prompts showed a significant effect of aspect by items, and a marginal effect by participants, when continuations were restricted to name-initial ones.
the whole event is atelic. If this is the case, an occasion relation may be expected to produce similar results in the two perfective conditions, with a singular vs. plural entity for the entity of transfer. But, to repeat, I consider that the referent bias may be what leads to discourse coherence rather than what is derived from discourse coherence.

Yet another factor that could affect the referent bias is co-locatedness of Source and Goal referents. Sub-analyses on co-locatedness of Source and Goal referents conformed to results shown by Rohde, Kehler, and Elman (2006), in that less co-located items increased Source bias more than co-located items. Rohde, Kehler, and Elman saw a meaningful increase of Source bias over Goal bias in the perfective condition with non-co-located verbs. However, in the present study’s experiments, the Source bias increased in both the imperfective-singular and the perfective-singular conditions in less-co-located items compared to co-located items. Interestingly, the perfective-plural condition showed no difference due to co-locatedness. This can be explained by the iterative nature of the events expressed in the perfective-plural condition, where the Source and Goal referents were both repeatedly accessible, and this was true in both co-located items and less-co-located items. As Rohde, Kehler, and Elman’s comparison was within the verb class of co-locatedness, and my comparison was between the verb classes, it is not possible to directly compare the increase of the Source bias in the imperfective condition in the two studies. But in their study, the Source bias in the perfective condition in co-located verbs (Class 1, co-located and guaranteed transfer, and Class 2, co-located and no guaranteed transfer) was around 50% in Class 1 and less than 40% in Class 2, while the Source bias in their Class 3 (non-co-located item, no guaranteed transfer) was over 60%, conforming to the observations in my study.
For the reasons I have discussed in this section, I conclude that this study’s results show that the perfective grammatical aspect invariably increases Goal bias, as opposed to the imperfective grammatical aspect, and that constraints on comprehenders’ construction of a situation model are not affected by lexical aspect information.

These conclusions support the claims on aspectual coercion. The previous studies that have explored aspectual coercion have treated it as a special case. In *The light flashed until dawn*, no individual lexical item encodes the iterative interpretation of the situation. This type of semantic computation is rare and special. The present study, however, indicates that such aspectual coercion is actually more general and ubiquitous than previously considered; as in the case of *The light flashed until dawn*, a mismatched lexical-grammatical aspect sentence like *Susan kicked balls to John* does not encode the end point of the event in *to kick balls*. Nevertheless, the results showed that comprehenders had a Goal bias, which indicated that comprehenders interpreted the situation with the end-state, i.e., as a completed event.

Finally, coercion is reported to be costly in sentence processing (Piñango et al., 1999, 2006, etc.). In this regard, the coercion hypothesis seems suitable to explain response latencies reported in previous relevant studies. For example, coercion may explain the results in Yap et al.’s (2009) study, which had accomplishment verb conditions—that is, accomplishments in both imperfective (mismatch) and perfective (match) conditions. For the reaction times (RTs) for picture selection between a matching picture (completed) and a mismatching picture (on-going), they found a marginal difference, while for the activity verb in the imperfective (match) vs. perfective (mismatch) conditions, the matching (on-going) condition had significantly faster RTs.
than the mismatching (completed) condition. This may be explained because no cost of coercion is expected with accomplishment verbs, which contain readily available sub-events of both on-goingness and end-state, while with the activity verb the comprehenders are coerced to create an interpretation that includes completion of the event. Similarly, it may explain why Madden and Zwaan (2003), who used all accomplishment verbs when testing imperfective vs. perfective effects, did not see a difference in RTs for picture recognition between on-going vs. completed pictures with the imperfective sentences while they did see a difference with the perfective sentences. Because in accomplishment verbs there are sub-events of both on-goingness and end-state, both are readily available and no cost of coercion is expected for the imperfective. I suggest that the present results, along with the results of these previous studies, indicate that there is coercion in the activity-perfective combination, and the latency observed in previous studies was not due to participants’ reluctance or the impossibility of choosing the interpretation of event completedness due to the unavailability of such a choice, but rather, due to aspectual coercion.

This idea would predict that, in the present study, the selection of a source referent would be slowed down in the perfective-plural condition compared to the perfective-singular condition. The present results showed a suggestion of this idea: The reaction times in Experiment A to responses choosing the Source referent were marginally slower following the perfective-plural condition (by participant, $M = 1838.80$, $SD = 565.71$) than following the perfective-singular condition (by participant, $M = 1751.74$, $SD = 559.87$) only by participants, but not by items ($t_1(35) = -1.54$, $p = 0.13$; $t_2(59) = -0.47$, $p = 0.64$, two-tailed). The opposite tendency appeared in Experiment B,
where the reaction times for Source referents were numerically faster following the perfective-plural condition (by participant, $M = 1395.92, SD = 277.14$) than following the perfective-singular condition (by participant, $M = 1497.17, SD = 473.02$) ($t_{1}(31) = 1.56, p = 0.13$; $t_{2}(58) = 1.29, p = 0.20$, two-tailed). These opposite tendencies may be because Experiment B used only pronouns for referents, and there are those who claim (as we have seen above) that reference to the subject (i.e., Source referent) is preferred in pronoun resolution whereas the repeated name penalty (as in Experiment A, which used proper names for referents for more than half the items) would bias responses to the Goal.

This study’s results thus, overall, support the coercion hypothesis, indicating that grammatical aspect plays a dominant role in interpreting an event situation, and in a mismatched combination of lexical and grammatical aspect, comprehenders are coerced to interpret that event (which has no end-point) as completed by the use of perfective grammatical aspect. There have been no other previous studies that explicitly showed this, to my knowledge.
One recent study detected a greater processing cost in reading imperfective sentences than perfective sentences. In Coll-Florit and Gennari (2011, testing in Spanish), imperfective sentences took readers longer to process than perfective sentences. However, a question remains whether the costs originate in grammatical aspect or in lexical aspect, or even in their interaction. Coll-Florit and Gennari claimed that the imperfective sentences are costly since they describe durative events while the perfective describes punctual events (Let us call this claim the Uniformity Hypothesis 1, as it assumes a uniform role for grammatical aspect in imperfective sentences). Coll-Florit and Gennari explain that durative events are costly as they elicit semantically more diverse associations, reflecting our experience, than non-durative events. But I propose that such costs may indicate a reader’s uncertainty about the duration of the described event in the imperfective sentences (Uniformity Hypothesis 2: Uncertainty Hypothesis). Another approach to this question is to ask whether they are the costs of comprehending the event itself. Since verbs and grammatical aspect were presented together in Coll-Florit and Gennari’s study, any processing costs of events were mixed with any costs of grammatical aspect. But other studies have shown that comprehenders make more detailed simulations of events in their minds when an event is described with the imperfective as opposed to the perfective (e.g., Bergen & Wheeler, 2010; Anderson, Matlock, & Spivey, 2010; Ferretti et al., 2007, among others). Considering this, it may be
only for the imperfective sentences that comprehenders’ reading times will reflect the
duration of an event described as short or long, while such difference of duration will not
affect the reading time results from perfective sentences (Interaction Hypothesis).

Yet another approach involves the processing costs related to highlighting parts of
an event. Chapter 3 discussed that the imperfective aspect highlights the ongoing nature
of an event and provides temporal information about the unfolding of events.
Surprisingly, a case that has not yet been investigated is when the imperfective aspect
highlights an event’s completion versus the ongoing nature of an event. To give a
Japanese example, the imperfective sentence John-ga (eki-ni) tsui-tei-ta means that John
had arrived (at the station), but not that John was arriving (at the station). This use of the
imperfective -tei- has a resultative reading that assumes the event’s completion as the
perfective does, and highlights the end-state. A question is whether this distinction in the
highlighting of parts of events affects the costs of reading the imperfective sentences
(Focusing Hypothesis). The perfective highlights the resulting end-state, and it has been
claimed to shut off access to the described process, while the use of the progressive, it has
been argued, accentuates the internal structure of an event (Carreiras, Carriedo, Alonso,
& Fernández, 1997; Magliano & Schleich, 2000; Madden & Zwaan, 2003; Ferretti, Kutas,
& McRae, 2007; Madden & Therriault, 2009. Also see Bergen & Wheeler, 2010); this
may explain why the costs differ between the two aspects.

In summary, much remains unclear about the roles of grammatical aspect and
lexical aspect in sentence processing. The following subsections summarize the four
hypotheses.
**Uniform Hypothesis 1 (Durativity Hypothesis).** Since grammatical aspect plays its role uniformly with any event, the imperfective sentences simply will have longer reading times than the perfective sentences, and this would not vary with short or long duration of the events. The assumption that supports this hypothesis is that the imperfective expresses durative events while the perfective expresses punctual events (Coll-Florit & Gennari, 2011).

**Uniform Hypothesis 2 (Uncertainty Hypothesis).** Because they lack information on event endedness, the imperfective sentences will make readers take longer than the perfective sentences, since the imperfective sentences leave readers uncertain about how long the event will last (it can even be infinite) while the perfective informs readers of an explicit boundary. If for any reason readers are informed of the event duration, the imperfective sentences would then become no more costly than the perfective sentences, since readers would know how long the event will continue.

**Interaction Hypothesis.** Since comprehenders make more detailed simulations of events in their minds when the event is described with the imperfective as opposed to the perfective, only for the imperfective sentences will comprehenders’ reading times reflect the duration of an event described as short or long, while such difference of duration will not show up in the reading time results with the perfective sentences.

**Focusing Hypothesis.** Since, in imperfective sentences, the processing costs caused by any of the mechanisms that underlie the other hypotheses relate to the ongoing nature of
an event versus its completion, the costs are observed when the imperfective sentences highlight the ongoing nature of an event, but not when they highlight the event’s completion. To restate this hypothesis in terms of the case of the Japanese imperfective -tei-, the costs are observed when the imperfective sentences express the progressive readings, but not when they express the resultative readings.19

This chapter presents four experiments (Experiments 2 to 5) to explore how grammatical aspect impacts the cost of processing in relation to event-related variables. The duration of events (short or long) and grammatical aspect (imperfective vs. perfective) are manipulated in all experiments. The relative timing of information available is manipulated by using English (Experiment 2) and Japanese (Experiments 3 and 4); readers encounter first grammatical aspect and then event duration in English sentences, while this order is reversed in Japanese sentences. The Japanese experiments also introduce another variable of which part of the event is highlighted, as a between-experiment factor. The difference in which part of an event is highlighted in sentences using identical forms of grammatical aspect is possible when using the Japanese imperfective marker -tei-. The imperfective -tei- can lead to the progressive reading (Experiment 3) or the resultative reading (Experiment 4) depending on verb types. An additional experiment using Japanese (Experiment 5) removes the variable of event

19 To test this hypothesis, all test items used in the present studies were manipulated so that -tei- sentences always had an unambiguous reading of either the progressive or the resultative. Moreover, Japanese is a verb-final language, and reading times at the region of the verb with grammatical aspect (Region 5) are related to the global readings, should sentences be locally ambiguous. Hence any differences of behaviors under this hypothesis will be observed in reading times at Region 5.
duration to provide a baseline for comparison, in which participants encounter grammatical aspect without encountering information about event duration.

Event duration is manipulated to be short or long by including a numerical expression, such as 3 holes vs. 18 holes of golf, or swim for 5 minutes vs. 50 minutes. An advantage of manipulating event duration with numerals is that it allows a within-item design, with short or long versions of otherwise identical events, rather than a between-item design that manipulates lexical items of verbs. Additionally, it avoids arbitrary conceptions of duration that may differ depending on readers or contexts. For example, an event intended to appear short, as expressed by for a while, could last 5 minutes according to one reader or in one context, or 15 minutes for another reader or in another context, and this variation would hinder comparison within and between participants and items. However, event duration described by numerical expressions is fixed and identical across participants and items.

Since imperfective sentences help readers create more detailed simulations of described events, I predict longer reading times in the imperfective than the perfective, and that a difference in reading times will correlate with short or long event duration with the imperfective sentences. No such difference is expected within the perfective sentences, as the perfective has been claimed to encapsulate or shut off access to the described process. The predictions will further depend on languages or word order, and changes in highlighting of event parts.

Section 5.1 presents Experiment 2 (English), which explored the case where event duration is read after grammatical aspect. The word order and tense-aspect morphology of English is similar to that of Spanish (the language tested in Coll-Florit & Gennari,
I predict a replication of Coll-Florit and Gennari’s results, where the imperfective leads to longer reading times than the perfective, but there should also be an interaction effect of event duration and grammatical aspect as soon as readers obtain the information on duration.

Section 5.2 presents Experiments 3 and 4 (Japanese), which tested cases where information on the duration of the event appears before grammatical aspect. Since comprehenders obtain the information on event duration before they encounter grammatical aspect, and if this helps readers know how long the event will continue, reading times may not be different between the imperfective and perfective sentences, if the costs seen in the imperfective sentences in Coll-Florit and Gennari are actually the costs of comprehenders’ uncertainty on how long the event will take. If this is not true, then, as in Experiment 2 (English), the imperfective sentences would show longer reading times. As readers are informed of event duration before grammatical aspect, I predict that any interaction effect of event duration and grammatical aspect will take place at the region of grammatical aspect, unlike in Experiment 2 (English).

In addition, Experiments 3 and 4 compare cases that differ by which part of the events are highlighted. Experiment 3 tests cases where the imperfective -tei- highlights the ongoing nature of an event by using verbs that induce a progressive reading, while Experiment 4 tests cases where the -tei- highlights the completion of an event by using verbs that induce a resultative reading. If the imperfective plays a role because it provides temporal information about the unfolding of events, the costs will be demonstrated only in Experiment 3, with no effect appearing in Experiment 4. If the highlighting of different
parts does not affect the cost of reading sentences with different aspects, there should be no difference between the results of Experiments 3 and 4.

Section 5.3 presents Experiment 5 (Japanese), which tested a case where no information on event duration is encountered before grammatical aspect. In this experiment, alone among Experiments 3–5, the Uncertainty Hypothesis will be supported if we observe longer reading times in the imperfective than the perfective, since readers will be uncertain about how long the event will last.

5.1 Event duration read after grammatical aspect (English)

5.1.1 Participants
Sixty-three students at the University of Hawai‘i participated in an about 15-minute long experiment for the compensation of course credit or a small bag of snacks. All were native speakers of English and some also spoke other language(s) natively (44%). The dominance of English in the non-monolingual speakers was self-reported as 50% or more.

5.1.2 Materials

Critical Sentences. The critical sentences consisted of 24 test items created by the researcher with the help of a native speaker of English. Each test item contained a short or long event in an imperfective or perfective sentence. All the sentences shared an identical structure for the first five phrases, i.e., the subject (Region 1), a verb with grammatical aspect (Region 2), event duration (Region 3), a modifying phrase (Region 4), and another adverbial (Region 5). These five fixed regions were followed by any optional
phrase(s), to have the sentence lengths within a range of five to nine phrases across test items. The critical regions are Regions 2 and 3, and additionally Region 4 for any spillover effects from Region 3. Table 5.1 illustrates the structure of the critical sentences.

Table 5.1. Structure of the critical sentences in Experiment 2

<table>
<thead>
<tr>
<th>Regions of Interest</th>
<th>Critical (V+Aspect)</th>
<th>Critical (Duration)</th>
<th>Post-Critical</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regions</td>
<td>Region 1</td>
<td>Region 2</td>
<td>Region 3</td>
</tr>
<tr>
<td>D.O. version</td>
<td>Tom</td>
<td>was golfing/</td>
<td>{18/three}</td>
</tr>
<tr>
<td></td>
<td></td>
<td>golfed</td>
<td>holes</td>
</tr>
<tr>
<td>Adv. version</td>
<td>Isaac</td>
<td>was sleeping/slept</td>
<td>for {12/two}</td>
</tr>
</tbody>
</table>

In Region 1, a proper name was used for the subject noun to represent a single character. Gender for the subject noun was counterbalanced such that female and male characters each appeared in half of the 24 items. Region 2 used either a past progressive (e.g., *was sleeping*) or simple past tense (e.g., *slept*) to manipulate grammatical aspect. Region 3 provided information on event duration.

Crucially, duration of event was manipulated by a numeral that appeared in Region 3. One of two forms was employed: either the direct object of the verb (D.O. version; e.g., *{18/three} holes*) or a temporal adverbial (Adv. version; e.g., *for {12/two} hours*). The reason for this variation was to avoid any effect of a particular linguistic form. Half of the 24 test items (12 items) employed a direct object while the other 12 items employed an adverbial phrase in Region 3. All of the adverbial items used the structure *for...* (e.g., *for x hours, days, minutes*, etc.). Consideration was given to how to present the numeral. For the sake of naturalness, the common practice was used in which numbers lower than ten were spelled out as words, and numbers of ten or higher were
given in Arabic numerals. Some items used *a* (or *one*), but most of the items used a number greater than one, so that plural morphology could appear in both the short and the long version of an expression.

Region 4 used a modifying phrase (e.g., *in his room*) as a post-critical region for any spillover effects from the preceding critical region, Region 3. Region 5 used another adverbial phase (e.g., *yesterday*) to avoid ending the sentence with Region 4. This was done to avoid mixing any effects in Region 4 with the wrap-up effects at a clause boundary, which have been reported to induce extra processing cost (Aaronson & Scarborough, 1976; Just & Carpenter, 1980; Rayner, Sereno, Morris, Schmauder, & Clifton, 1989). All region(s) in a sentence were given a reasonable and natural punctuation. The last region of each sentence included a period, which is considered one character when calculating residual reading times (Trueswell, Tanenhaus, & Garnsey, 1994).

In the sentences presented to each participant, a quarter had verb phrases with imperfective aspect with long duration (7), a quarter had perfective with long duration (8), a quarter had imperfective with short duration (9), and a quarter had perfective with short duration (10). See Appendix C for a complete list of critical sentences used in Experiment 2. The following example sentences illustrate the four conditions in an adverbial version.

(7) Condition 1. Imperfective-Long: Isaac/ was sleeping/ for 12 hours/ in his room/ yesterday./
(8) Condition 2. Perfective-Long: Isaac/ slept/ for 12 hours/ in his room/ yesterday./
(9) Condition 3. Imperfective-Short: Isaac/ was sleeping/ for two hours/ in his room/ yesterday./
(10) Condition 4. Perfective-Short: Isaac/ slept/ for two hours/ in his room/ yesterday./
**Filler Sentences.** The 48 filler sentences were created with a variety of sentence structures, including passives, a main clause with a subordinate clause, various tenses, plural and/or inanimate subject nouns, and so on. Filler sentences had a range of three to ten phrases. Gender for the subject noun was counterbalanced. Phrase division was made at reasonable and natural boundaries, though numbers of letters included in a region varied in a range from 1 (e.g., *I*) to 24 (e.g., *the Hawaii football team*).

**Confirmation Questions.** The 72 confirmation questions were created by the researcher with the help of a native speaker of English. These were simple yes/no questions asking about the content of each critical and filler sentence. These confirmation questions served to maintain and monitor participants’ attentiveness to the reading. Half of all the question sentences required “yes” and the other half required “no” for the correct answer.

**5.1.3 Procedure**

Participants read 24 critical and 48 filler sentences one phrase at a time, with the speed determined by their own key pressing, and then answered a simple yes/no question about each sentence. The 72 sentences were presented in a pseudo-random order, such that participants would see one critical sentence per two filler sentences. Each sentence was presented left-aligned, in a phrase-by-phrase moving-window presentation on the computer screen. The font size was selected to be small enough so that the first five regions of the 24 critical sentences would appear on a single line but large enough to maintain readability. Participants were instructed to read the segments by pressing the spacebar and to press the keys labeled “yes” or “no” to answer the questions. They were
instructed to read the sentences and answer the questions as quickly and accurately as possible, but to make sure to understand each word. Since the computer was measuring reading times, participants were also asked to keep both hands on the keyboard during the test session to avoid any latency in key pressing unrelated to reading. Each phrase was replaced by an underbar upon pressing of the space bar, and the last phrase of each sentence was immediately followed by a question sentence upon pressing of the spacebar. The question sentences were presented alone in the middle of the screen and in green to signal the transition from the reading sentences presented in black. The “yes” key was positioned on the right and the “no” key on the left of the key board. The question sentence stayed on the screen until the participant pressed either the “yes” or the “no” key. Then a blank screen with a fixation cue in the middle appeared until participants pressed the spacebar to move on to the next sentence. A practice session with four items preceded the main session. The instructions to participants were given visually and orally: Participants were instructed orally while seeing screens that presented written instructions. Presentation of stimuli was in pseudo-random order across items and participants, and measurements were controlled and collected by E-prime software 1.2.

5.1.4 Conditions

Experiment 2 tested four conditions with two variables: (i) grammatical aspect with two levels (Imperfective, Perfective), and (ii) event duration of a described event with two levels (Long duration, Short duration). Reading times at the sentence positions of Region 2 (verb plus grammatical aspect), Region 3 (duration phrase), and Region 4 (post-duration phrase for any spillover effect) were the dependent measures in this experiment.
5.1.5 Predictions

Experiment 2 (English) tested the three particular hypotheses in which the on-line sentence processing of events may, or may not, be impacted by grammatical aspect alone, or by an interaction of event information and grammatical aspect. Since English is similar to Spanish in terms of word order and structural complexity, I predicted that this experiment would replicate the results of Coll-Florit and Gennari (2011) up to the position of the verb with grammatical aspect (Region 2) and through the following region (Region 3) for spillover effects. I also predicted that the grammatical aspect would impact the processing of the test sentences once readers encountered the information on event duration (Region 3). Specific predictions related to each hypothesis follow.²⁰

**Uniform Hypothesis 1 (Durativity Hypothesis).** Since grammatical aspect plays its role uniformly for any event, the imperfective sentences should simply have longer reading times than the perfective sentences, and this should not vary by short or long duration of the events. Replicating Coll-Florit and Gennari’s (2011) results, the region for the verb with grammatical aspect (Region 2) and the successive region (Region 3) for spillover effects will show longer reading times in the imperfective sentences than in the perfective sentences.

**Uniform Hypothesis 2 (Uncertainty Hypothesis).** Since participants read information about the grammatical aspect (at Region 2) before they read information about the event

---
²⁰ The *Focusing Hypothesis* does not apply to Experiment 2 (English).
duration (at Region 3), the region for the verb with grammatical aspect (Region 2) will show longer reading times in the imperfective sentences than in the perfective sentences. In Region 3, however, as readers obtain the information on event duration, reading times may not increase as much in the imperfective sentences as observed in Coll-Florit and Gennari’s (2011) study, since at Region 3, readers are certain about how the described event continues and are no longer uncertain about the event’s endedness.

**Interaction Hypothesis.** Since comprehenders make more detailed simulations of events in their minds with imperfective sentences than with perfective sentences, only for the imperfective sentences will comprehenders’ reading times reflect the duration of an event described as short or long, while such difference of duration will not affect the reading times for the perfective sentences. Hence, reading times will increase at Region 3 (and possibly at Region 4 for spillover effects) only in the imperfective sentences. In other words, I expect an interaction effect of grammatical aspect and event duration at Regions 3 and 4.

### 5.1.6 Data analysis

**Selected Participants and Items.** First, the participants’ accuracy rate on the confirmation questions was calculated based on the entire set of trials (including critical and filler items), with the idea that participants should be excluded if their accuracy rate was below 80%. No participants were removed due to inaccuracy (the average of the accuracy rates from 64 people was 95.7%). Second, the items’ accuracy rates on the confirmation questions were also calculated based on responses from all participants, with the idea that
critical items should be excluded if their accuracy rate was below 80%. No critical items were removed due to this criterion (the average accuracy on the 24 critical items was 95.1%). All incorrect responses were filtered out for subsequent data cleaning procedures.

Third, outlying reading times were removed, with a fixed cutoff time of 3000 milliseconds, and the outliers were replaced with the value of 3000 milliseconds (70 reading times, i.e., 0.28% of reading times, were removed and replaced), with the idea that outlying times would skew the mean values and standard deviations in subsequent procedures. Fourth, outlying participants were searched for by calculating the mean value and standard deviation of the reading times of the entire set of trials (including critical and filler probes) for each participant, with the idea that participants should be excluded if their mean reading time was below or above 3 standard deviations of the overall mean. One participant\(^{21}\) (Mean = 1118 milliseconds) was excluded due to this criterion (the average of overall mean reading times from 64 people was 660 milliseconds). Fifth, outlying items were searched for by calculating the mean value and standard deviation of reading times on critical items from all participants, with the idea that items should be excluded if their mean reading time was below or above 3 standard deviations of the overall mean. No items were removed due to this criterion (the average of overall mean reading times from the 24 critical sentences was 676 milliseconds). Finally, in order to have an equal number of participants for each condition, three more participants (those with the lowest accuracy rates on the confirmation questions) were removed. As a result, data from 60 people (i.e., 15 participants for each of four conditions) were left for the subsequent analyses.

\(^{21}\) The data patterns are similar even when this participant’s data are retained in the analyses.
**Selected Response Times.** The mean value and standard deviation of the reading times (including critical and filler items) were calculated for each participant. Trials were replaced if their reading times were below or above 3 standard deviations of the value for each participant. With this criterion, 2.05% of overall raw reading times on the entire set of trials (including critical and filler probes) were replaced with either 3000 milliseconds or 3 standard deviations above or below the participant’s mean, whichever was applicable.

**Residual Reading Times.** I applied a standard treatment for calculating residual reading times (e.g., Trueswell et al., 1994; See also Frazier and Rayner, 1982; Altmann, Garnham, and Dennis, 1992; Ferreira and Clifton, 1986; Rayner, Carlson, and Frazier, 1983; Rayner, Garrod, and Perfetti, 1992). The conversion from raw (cleaned) reading times to residual reading times was essential to adjust for differences in word length across conditions as well as overall differences in participants’ reading rates (see Trueswell et al., 1994 for discussion). For each participant, a regression equation was constructed from word length using reading times from filler and experimental items. However, reading times at critical regions (i.e., Regions 2, 3, and 4) of the critical items were set aside because these reading times were expected to be affected by experimental factors across conditions, and therefore they would skew predicted reading times based on word length. The reading times used were those that were followed by correct responses to the confirmation questions. Residual reading times were obtained by subtracting the expected reading times calculated by regression equations from the actual reading times by each participant.
The mean multiple R and the mean R square of regression equations across 60 people were 0.38 and 0.16, respectively. These values indicate that the strength of relation between word length and prenormalized reading times was, overall, 0.38 on a scale of 0 to 1 (with 1 the maximum), and that the equation can predict 16% of the estimated reading times.

5.1.7 Results

Mean phrase-by-phrase residual reading times computed across participants are plotted in Figure 5.1 for each condition in each of the five regions. Figure 5.2 presents mean phrase-by-phrase raw reading times computed similarly across participants for each condition in each of the five regions. Although not all of the data patterns for the residual times are numerically the same as those for the raw times, the critical tests on both data sets showed the same pattern, except for Region 2. Not all of the critical tests reached significance in the raw data. See Table 5.2 for a complete set of raw and residual reading time means by condition and position. In what follows, I primarily report the statistical results from residual reading times, but I also refer to the results from raw reading times.

No tests are reported for Region 1 (the subject), as the conditions did not differ at this point (for both residual and raw reading times [RTs], all Fs < 1). The residual reading times in the first critical region, Region 2 (the verb with grammatical aspect), show that the imperfective conditions were read significantly faster (mean by participant = -113.42,
Figure 5.1. Mean residual reading times in each region from Experiment 2, by participants, in milliseconds, and 95% confidence intervals for each condition group

Figure 5.2. Mean raw reading times in each region from Experiment 2, by participants, in milliseconds, and 95% confidence intervals for each condition group

Isaac was sleeping slept for 12 hours two hours in his room yesterday.

Tom was golfing golfed 18 holes three holes with his colleagues last Sunday.
Table 5.2. Raw (and residual) reading times per word in each region from Experiment 2, by participant, in milliseconds

<table>
<thead>
<tr>
<th>Condition</th>
<th>Region1</th>
<th>Region2</th>
<th>Region3</th>
<th>Region4</th>
<th>Region5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Imp - Long</td>
<td>691(111.6)</td>
<td>597(-106.6)</td>
<td>681(-20.0)</td>
<td>737(-12.4)</td>
<td>738(22.9)</td>
</tr>
<tr>
<td>Perf - Long</td>
<td>688(112.1)</td>
<td>534(-43.4)</td>
<td>639(-63.8)</td>
<td>708(-40.2)</td>
<td>760(42.7)</td>
</tr>
<tr>
<td>Imp - Short</td>
<td>692(115.0)</td>
<td>584(-120.3)</td>
<td>650(-66.8)</td>
<td>719(-30.6)</td>
<td>779(59.1)</td>
</tr>
<tr>
<td>Perf - Short</td>
<td>695(116.9)</td>
<td>520(-56.4)</td>
<td>617(-101.0)</td>
<td>722(-19.6)</td>
<td>744(15.7)</td>
</tr>
</tbody>
</table>

*Abbreviations: Imp = Imperfective, Perf = Perfective

$SE$ [standard error]= 13.49) than the perfective conditions (mean by participant= -49.89, $SE=11.01$) both by participants ($F_1(1, 59) = 28.73, p<.0001$) and by items ($F_2(1, 23) = 28.18, p<.0001$). However, the raw reading times showed the opposite effect: The imperfective conditions were read significantly more *slowly* (mean by participant= 590.35, $SE=18.18$) than the perfective conditions (mean by participant= 526.65, $SE=16.35$), both by participants ($F_1(1, 59) = 30.22, p=.0033$) and by items ($F_2(1, 23) = 24.41, p<.0001$), as predicted for processing differences between the two conditions by Uniform Hypotheses 1 and 2. These contrasting results between residual and raw times were only observed in Region 2. There was no main effect of event duration (residual RTs: both $F$s $<1.6$; raw RTs: both $F$s $<1.7$, by participants and by items), or interaction between aspect and event duration (residual and raw RTs: both $F$s $<1$ by participants and by items) in this region.

In the second critical region, Region 3 (the event duration phrase), the residual reading times show that the imperfective conditions were read significantly slower (mean by participant= -43.38, $SE=15.90$) than the perfective conditions (mean by participant= -82.42, $SE=13.83$), both by participants ($F_1(1, 59) = 9.40, p=.0033$) and by items ($F_2(1, 23) = 7.198, p=.013$). Unlike the results of the raw reading times in Region 2, which
contrasted with those for the residual reading times, the results for raw reading times in Region 3 were in the same direction as those for residual reading times, both by participants ($F_1(1, 59) = 9.12, p = .0037$) and by items $F_2(1, 23) = 7.26, p = .013$). In addition, the long duration conditions were read significantly more slowly (mean by participant= -41.91, $SE = 14.77$) than the short duration conditions (mean by participants= -83.90, $SE = 17.18$), both by participants ($F_1(1, 59) = 5.87, p = .019$) and by items $F_2(1, 23) = 7.15, p = .014$) in this region. However, the raw reading times showed only a suggestion of a main effect of event duration, both by participants ($F_1(1, 59) = 2.69, p = .106$) and by items ($F_2(1, 23) = 2.68, p = .115$). Contrary to the Interaction Hypothesis, there was no interaction between aspect and event duration (residual and raw RTs: both $Fs < 1$ by participants and by items) in this region.

In the last critical region, Region 4 (the post event duration phrase for any spillover effects), there were no main effects of aspect or event duration (residual and raw RTs: all $Fs < 1$ by participants and by items). There was a suggestion of interaction between aspect and event duration, but the effect only nearly reached significance by items ($F_2 (1, 23) = 3.89, p = .061$) and was not significant by participants ($F_1 (1, 59) = 2.02, p = .16$) in residual times. Raw times failed to show an interaction between aspect and event duration either by items ($F_2 (1, 23) = 2.195, p = .15$) or by participants ($F_1 (1, 59) = 1.13, p = .29$). Paired $t$-tests on the interaction effect across all the possible pairs in the four conditions revealed no significant differences. The largest suggestion was that, in long event conditions, the imperfective was read numerically more slowly (mean by participants = -12.36, $SD = 117.25$) than the perfective (mean by participants = -40.17,
The small hint of interaction between aspect and event duration continued in the successive region, Region 5, where it was marginal by items ($F_2 (1, 23) = 3.62, p = 0.0696$) and not significant by participants ($F_1 (1, 59) = 2.45, p = .12$) in residual reading times, and not significant in raw times ($F_1 (1, 59) = 1.95, p = .168; F_2 (1, 23) = 2.66, p = .12$). Paired t-tests on the interaction effect across all the possible pairs in the four conditions revealed only weak patterns without significance: (a) There was a suggestion that long event sentences (mean by participants = 22.88, $SD = 121.18$) were read faster than short event sentences (mean by participants = 59.09, $SD = 161.48$) in imperfective conditions, both in the raw and residual reading time analyses, although the effects were only marginal by items and did not reach significance by participants (residual RTs: $t_{1}(59) = -1.37, p = 0.18; t_{2}(23) = -1.98, p = 0.060$; raw RTs: $t_{1}(59) = -1.46, p = 0.15, t_{2}(23) = -1.97, p = 0.061$), and (b) there was a numerical pattern that, in short event conditions, the imperfective was read more slowly (mean by participants = 59.09, $SD = 161.48$) than the perfective (mean by participants = 15.71, $SD = 149.76$) (residual RTs: $t_{1} (59) = 1.69, p = 0.096; t_{2} (23) = 1.63, p = 0.12$; raw RTs: $t_{1} (59) = 1.16, p = 0.25, t_{2} (23) = -1.24, p = 0.23$).

No other comparison in the t-tests showed a significant difference. There were no main effects of aspect or event duration (residual and raw RTs: all $Fs < 1$ by participants and by items) in this region.
5.1.8 Discussion

In this section, I follow a standard of basing my conclusions on the cases in which the results for the raw and residual RTs agree. The only place in which the results showed significance for both raw and residual RTs was in Region 3, at the post verb-with-aspect region: The imperfective sentences were read more slowly than the perfective sentences, as similarly observed by Coll-Florit and Gennari (2011). The remainder of this section provides details about the rest of the analyses, region by region. It is crucial to the present study that the results revealed no interaction of aspect and event duration in Regions 3, 4, and 5. At Region 2 (the verb with the grammatical aspect), a significant main effect of aspect appeared for both the raw and residual RTs, but in opposite directions. I will also present my interpretation of this discrepancy in this section.

At Region 3, the imperfective conditions were read more slowly than the perfective conditions in both raw and residual reading times. Region 3 presented an event duration expression, yet, crucially, a main effect of aspect was confirmed in this region.

The main effect of aspect at the post-verb position was also observed by Coll-Florit and Gennari (2011), along with a main effect of aspect at the preceding region, the region of the verb with grammatical aspect. In Coll-Florit and Gennari’s study, they did not see the effect in the post-verb region when the context was provided preceding the critical sentence, while they did observe it when the aspect sentence was presented alone without a preceding context sentence. Coll-Florit and Gennari explain that the narrative contexts that preceded the test sentence allowed readers to anticipate discourse entities, or to form situation models that already contained them. In the absence of context, new entities (i.e., the direct object) must be introduced into the situation being built, and their
meanings must be integrated with that of the verb for the construction of the durative or non-durative situation, hence the effect at the post-aspect region in the non-context experiment. One must keep in mind, in the present study, (i) that in half of the critical sentences, the material at Region 3 was not the direct object but an adverbial phrase, which was not an argument but an adjunct (i.e., optional to the sentence), and (ii) that the lexical items were different in this region (i.e., short or long event duration) due to manipulation of event duration, unlike in Coll-Florit and Gennari’s study. Despite these differences, the present study showed the main effect of aspect in the post-verb region, at Region 3. Hence, the present results may be taken as evidence of a continuing effect of aspect from the previous region, rather than of effects of informational cost, as discussed by Coll-Florit and Gennari.

The main effect of aspect in Region 2 revealed significance in both the raw and residual reading times, but it was in different directions. The results of the raw reading times suggested that the imperfective conditions were read more slowly than the perfective conditions, while the results of the residual reading times suggested that the imperfective conditions were read faster than the perfective conditions. However, this difference is probably due to a failure in estimating reading times with the regression equations. One must keep in mind that the mean R-Squared across participants was 16% for the equations. This means that it is possible to perfectly predict only 16% of the reading times from the word lengths. Furthermore, in the present study, the nature of residual times in the imperfective conditions may be less reliable than in the perfective conditions. For instance, the imperfective condition presented “was + [space] + V-ing” at Region 2, while the perfective presented “V-ed”; the former contained two words with a
space in between, while the latter contained only one word. Therefore, a likely possibility is that because of an overestimate on reading times, too much was subtracted from the base times, such that the residual times became too short in the imperfective, leading to the present conflict in results. Furthermore, it is important to note that the effects of aspect were significant in the following region, Region 3, in the same direction for both raw and residual times, where the imperfective sentences were read more slowly than the perfective sentences. Hence, the main effect of aspect observed in Region 3 seems to reflect a spillover effect from the previous region, Region 2, as can often occur in self-paced reading. To conclude, I interpret the main effect of aspect observed at Region 2 to reveal that the imperfective sentences were read more slowly than the perfective sentences, conforming to both Uniform Hypotheses 1 and 2, although the inconsistent reading time data do not provide strong evidence for this position.

In Region 3, an interaction of aspect and event duration would have provided support for the Interaction Hypothesis, but there was no such interaction. Comparison within the imperfective conditions shows that the long events were processed numerically more slowly in Regions 3 and 4, but this difference seems to be a reflection of a main effect of event duration at Region 3, without statistically supported interaction. The main effect of event duration at Region 3, which presented an event duration expression, was significant in residual times and was marginal in raw reading times: Long event conditions were processed more slowly than short event conditions. However, the effect was not present in other regions, either prior to or following Region 3. Therefore, this effect of event duration seems more local than global. It is worth noting that although the word lengths were supposed to be different between the conditions, in Region 3 they
were rather similar in the short and long duration conditions because the number that was used to manipulate duration was an Arabic numeral for numbers of ten and more, but a word for numbers smaller than ten. Thus, a likely possibility is that the adjusted reading times might underestimate the time needed to process the numerals in the long conditions, meaning that the Arabic numerals took longer to read than the estimate for the letter length. That is, the fact that the observed effects are local may point to the main effect of duration being an artifact of processing time for the numbers that were presented differently between the conditions.

At Region 5, the data implied that (a) there was a suggestion of long event sentences being read faster than short event sentences in imperfective conditions (by items) and (b) that in short event conditions the imperfective was read more slowly than the perfective (by residual times), but neither of these suggestions of an interaction of aspect and event duration reached significance. Because the lexical items in Regions 4 and 5 were identical across the four conditions, any observed effects in these regions should reflect any effects from the preceding regions. These effects are, however, isolated from any spillover from the preceding regions, because there were no interaction effects prior to Region 5: No suggestion of interaction was observed in the prior region, Region 4, and, in particular, there was no interaction at Region 3, which was the region where aspect and event duration first became available to the readers. Thus, the Interaction Hypothesis does not gain strong support in the present study. Irrelevant to the hypothesis, the numerical interactions at Region 5 might reflect that, as readers took longer to read imperfective sentences, they needed more time to process long event sentences in the imperfective at Region 4. This would mean that readers took less time in long event
sentences than in short event sentences in the imperfective at Region5, and the reading times for short event sentences at Region 5 were relatively longer in the imperfective than in the perfective. One must keep in mind, though, that the effects did not quite reach significance.

In sum, among the three tested hypotheses, the results from Experiment 2 overall are compatible with Uniform Hypotheses 1 and 2 (with explainable inconsistency between raw and residual reading times in Region 2) but not with the Interaction Hypothesis. The result of a main effect of aspect at Region 3 might be taken as some evidence against the predictions of Uniform Hypothesis 2. Because this hypothesis posits that information on event duration (whether it is short or long) may help reduce processing costs in the imperfective sentences by allowing readers to know how long the event lasts, the reading times at Region 3 would be predicted to decrease in the imperfective sentences and reduce the difference between the aspect conditions. However, Region 3 showed a main effect of aspect without differences between the imperfective and perfective conditions. This possibility that Uniform Hypothesis 2 is more viable than Uniform Hypothesis 1 was explicitly questioned and tested in experiments using Japanese sentences, which are presented in the following sections.

5.2 Event duration read before grammatical aspect (Japanese)

Experiments 3 and 4 were identical, with the exception that Experiment 3 focused on, and was conducted with, verbs that induce progressive readings in the imperfective sentences (the Progressive study), and Experiment 4 used verbs that induce resultative readings in the imperfective sentences (the Resultative study).
5.2.1 Participants (Progressive, Resultative)

Forty undergraduate students at Tohoku University (in Sendai, Japan) participated in the Progressive study (Experiment 3), and a separate group of 20 undergraduate students at Tohoku University participated in the Resultative study (Experiment 4), for compensation of $7 (approximately 500 yen) or $12 (approximately 800 yen). Participants were compensated $7 when they participated in this study alone, or $12 when they also participated in another experiment (Experiment 6 or 7) afterward. All were self-reported to be monolingual native speakers of Japanese.

5.2.2 Materials (Progressive, Resultative)

Experiments 3 and 4 used materials that were as similar as possible to those used in the English version (Experiment 2). The materials for Experiments 3 and 4 were identical except for slight differences that resulted from the nature of the verbs that they used. Tables 5.3 and 5.4 illustrate the structure of the critical sentences in Experiments 3 and 4, respectively. The differences consist in one versus two variations in structure, the nouns that are used for the subject (Region 1), and the forms of the temporal adverbial that appears in Region 3.

**Critical Sentences.** The 24 critical sentences were created by the researcher using verbs that induce progressive readings in the imperfective sentences for the Progressive study (Experiment 3), and verbs that induce resultative readings in the imperfective sentences for the Resultative study (Experiment 4). Each test item contained a short or long event
Table 5.3. Structure of the critical sentences in Experiment 3 (Progressive)

<table>
<thead>
<tr>
<th>Regions of Interest</th>
<th>Critical (Duration)</th>
<th>Post-critical</th>
<th>Critical (V+Aspect)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regions</td>
<td>Region 1</td>
<td>Region 2</td>
<td>Region 3</td>
</tr>
<tr>
<td>D.O. version</td>
<td>成美はNarumi-Top</td>
<td>今朝this morning</td>
<td>{80/12}枚の皿を</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>{80/12}-classifier-Gen-plates-Acc</td>
</tr>
<tr>
<td>Adv. version</td>
<td>淳はAtsushi-Top</td>
<td>昨日yesterday</td>
<td>{12/2}時間for {12/2} hours</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5.4. Structure of the critical sentences in Experiment 4 (Resultative)

<table>
<thead>
<tr>
<th>Regions of Interest</th>
<th>Critical (Duration)</th>
<th>Post-critical</th>
<th>Critical (V+Aspect)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regions</td>
<td>Region 1</td>
<td>Region 2</td>
<td>Region 3</td>
</tr>
<tr>
<td>Adv. version</td>
<td>美穂はMiho-Top</td>
<td>カバンにon her bag</td>
<td>{10/3}日間for about {10/3} days</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(explained below). All the sentences shared an identical structure with five fixed phrases: the subject (Region 1), a modifying phrase (Region 2), event duration (Region 3), another modifying phrase (Region 4), and a verb with grammatical aspect (Region 5). The critical regions are Regions 3 and 5, and, additionally, Region 4 for any spillover effects from Region 3.

24 Phrases were divided at reasonable boundaries rather than at bunsetsu in order to be consistent with the English versions. For example, Region 3 has 12枚の皿を‘twelve plates’ as one phrase, which could be two phrases, 12枚 and 皿を, if punctuated by bunsetsu.
In Region 1, Experiment 3 (Progressive) used a proper name with a topic marker -wa for the subject noun to represent a single character. Experiment 4 (Resultative) used either a proper name or a common noun for an inanimate object, with a topic marker -wa, as the subject noun in the resultative sentences is often a patient or theme that undergoes a change of state. The topic marker -wa was chosen over the nominative marker -ga for naturalness, and considering that no difference in effects between these markers has been reported in sentence processing (Ueno & Kehler, 2010). Gender for the subject noun was counterbalanced such that female and male characters each appeared in half of the stimuli. Region 2 contained an adverbial phase (e.g., yesterday) so that all the critical sentences would be naturally expressed within the five fixed phrases.

Crucially, duration of event was manipulated by a numeral that appeared in Region 3. One of two forms was employed for this in Experiment 3: the direct object of the verb (D.O. version; e.g., \{80/12\}枚の皿を ‘\{18/12\} plates’, or a temporal adverbial (Adv. version; e.g., 12時間/二時間 ‘for \{12/2\} hours’). The reason for this variation in structure was to avoid any effects from the particular linguistic form. Half of the 24 test items (12 items) were created in a D.O. version while the other 12 items were in an adverbial version. All items of the D.O. version used the structure -の…を (-no…o) ‘-Gen…Acc’ as the standard form for the expression. All items in the adverbial version used the structure of x間 (x-kan) ‘for x’ (e.g., x時間 ‘for x hours’; x分間 ‘for x minutes’; x日間 ‘for x days’, and so on). Experiment 4 employed only the adverbial versions of the 24 test items, because duration adverbials can modify the duration of the result state.

---

25 An alternative form could be …を - (…o-) ‘…Acc-’ (e.g., 皿を {80/12}枚 ‘plates-Acc-numeral-classifier’).
while the duration described by the numerals in the D.O. would modify the duration of the action that leads to the resultative state, but not the duration of the result state. One of two forms was employed for this expression in Experiment 4: \( x \)の間 (\( x -\text{no-aida} \)) ‘for the duration of/during the period of \( x \)’ or \( x \)間ほど (\( x -\text{kan-hodo} \)) ‘for about \( x \)’, because these forms of expression for the resultative sentences sound more natural than \( x \)間 (\( x -\text{kan} \)) ‘for \( x \)’. Consideration was given to how to present the numeral. For the sake of naturalness based on common usage, numbers lower than ten were given in kanji (Chinese characters), and numbers of ten and higher were given in Arabic numerals.

For D.O. sentences, Region 4 presented an adverbial phrase (e.g. in his room) as a post-critical region for any spillover effects from the preceding critical region, Region 3. For transitive Adverbial sentences, Region 4 presented the D.O.; otherwise it presented another adverbial phrase. Region 5 used either a past progressive (V-\( \text{tei-ta} \)) or simple past tense (V-\( \text{ta} \)) for manipulation of grammatical aspect.

Of the sentences that each participant saw, a quarter had verb phrases with imperfective aspect with long duration (11), a quarter had perfective with long duration (12), a quarter had imperfective with short duration (13), and a quarter had perfective with short duration (14). See Appendixes D and E for a complete list of critical sentences used in Experiments 3 and 4, respectively. The following example sentences illustrate the four conditions.

(11) Condition 1. Imperfective-Long:淳は/ 昨日/ 12 時間/ 彼の部屋で/ 眠っていた。
(13) Condition 3. Imperfective-Short:
淳は/ 昨日/ 二時間/ 彼の部屋で/眠っていた。
(14) Condition 4. Perfective-Short:
淳は/ 昨日/ 二時間/ 彼の部屋で/眠った。
Gloss: Atsushi-Top/ yesterday/ for {12/2} hours/ in his room/ {sleep-tei-ta/ sleep-ta}./
English Translation: ‘Atsushi was sleeping/slept in his room for 12/2 hours yesterday.’

**Filler Sentences.** The 48 filler sentences used in the Progressive study (Experiment 3) and the Resultative study (Experiment 4) were identical. The sentences were created by the researcher to have a variety of sentence structures, including passives, a main clause with a subordinate clause, various tenses, plural and/or inanimate subject nouns, and so on. Gender for the subject noun was counterbalanced. The filler sentences had two to six regions corresponding to phrases. Phrase division was made at reasonable and natural boundaries rather than at bunsetsu, and the number of letters included in a region varied from two (e.g., 道で‘on the road’) to 11 (e.g., フランス行きの飛行機に‘on an airplane bound for France’). The last region of each sentence included a period, which is considered one character when calculating residual reading times (Trueswell, Tanenhaus, & Garnsey, 1994).

**Confirmation Questions.** Each item in the experiment was followed by a confirmation question. The 24 confirmation questions for the critical sentences were created separately for the Progressive study (Experiment 3) and the Resultative study (Experiment 4), by the researcher. The 48 confirmation questions created for the filler sentences by the researcher were used in both Experiments 3 and 4. They were simple yes/no questions asking about the content of each critical and filler sentence. These confirmation questions
served to maintain and monitor participants’ attentiveness to the reading. Half of all the question sentences required “yes” and the other half required “no” for the correct answer.

5.2.3 Procedure (Progressive, Resultative)

The procedure used for the Progressive study (Experiment 3) and the Resultative study (Experiment 4) was nearly identical to that used for the English study (Experiment 2), with some minor adjustments to the studies conducted in Japanese.

As each new region appeared, the preceding phrase was replaced by a blank space rather than an underbar. This modification was made because in Japanese fonts the underbars appear as dotted lines, which look unnatural, rather than solid lines (as in the English study). The stimuli were presented in random order across items and participants, and measurements were controlled and collected by E-prime software 2.0, which allowed the Japanese fonts. “Yes” and “no” keys for answering the questions were positioned a little closer to the center of the keyboard than in the English study, reflecting feedback from participants in the English study, which was conducted earlier than the Japanese studies.

In the Japanese experiments, one of the critical regions is at the sentence-final position. To avoid participants’ staying on this final region for other reasons than reading, the experiments were implemented so that each phrase would automatically change to the next phrase after five seconds. Participants were informed of this time constraint prior to the main session. The reason behind this is because it was noticed that some participants spent a while at the final phrase before they moved on to a confirmation question during
the English study. This seemed to be the participants’ strategy to prepare to answer a question rather than being related to reading.

Except for these modifications, the procedure used for Experiments 3 and 4 was identical to that used in Experiment 2.

5.2.4 Conditions (Progressive, Resultative)

The four conditions (2 grammatical aspects by 2 event durations) were tested in Experiments 3 and 4 as in the English study (Experiment 2). The dependent measure was reading times at the sentence positions of Region 3 (duration phrase), Region 5 (verb plus grammatical aspect), and Region 4 (post-duration phrase for any spillover effect).

5.2.5 Predictions (Progressive, Resultative)

Experiments 3 and 4 tested the four particular hypotheses that the on-line sentence processing of events may, or may not, be impacted by grammatical aspect alone, or by an interaction of event information and grammatical aspect, or by a change in what part of the event is highlighted. Unlike Spanish or English, Japanese is a verb-final language. Since the information on the duration of events is available before grammatical aspect, I predicted that these experiments would not replicate the results of Coll-Florit and Gennari’s study (2011) at the position of the verb with grammatical aspect (Region 5) if the information on event duration helps readers to reduce processing costs that relate to uncertainty about how long the described event will take. Since information on duration of event is available before grammatical aspect, an interaction effect of event duration and grammatical aspect would be expected as soon as readers read the grammatical
aspect, unlike in Experiment 2 (English). In the Japanese studies, these hypotheses do not predict differences at the region for information on duration of events (Region 3), or the following region for any spillover effects (Region 4). Since information on the duration of the event is available before information on grammatical aspect, any effects of grammatical aspect (which may, or may not, interact with information on the duration of the event) should emerge as soon as readers read the grammatical aspect (Region 5), but not at Regions 3 or 4. Specific predictions follow.

**Uniform Hypothesis 1 (Durativity Hypothesis).** Since grammatical aspect plays its role uniformly with any event, and as the imperfective expresses durative events while the perfective expresses punctual events (Coll-Florit & Gennari, 2011), the imperfective sentences would have longer reading times than the perfective sentences without varying based on short or long duration of the events. Coll-Florit and Gennari explain that durative events are costly as they elicit semantically more diverse associations, reflecting our experience, than non-durative events. Reading times at the region for the verb with grammatical aspect (our Region 5) will be longer in the imperfective sentences than in the perfective sentences, replicating the results of Coll-Florit and Gennari, which revealed an increase of reading times at the region of the verb with grammatical aspect (and its successive region) in the imperfective sentences compared to the perfective sentences.

**Uniform Hypothesis 2 (Uncertainty Hypothesis).** If longer reading times for imperfectives in Coll-Florit and Gennari’s study (2011) were due to uncertainty about
how long the event would last, I predict that reading times for grammatical aspect (Region 5) will not be longer in the imperfective sentences than in the perfective sentences, in contrast to Coll-Florit and Gennari’s results. Since readers will have already become certain about the length of the described event at the region for event duration (Region 3), the readers will no longer be uncertain about the event’s endedness, which could otherwise cause processing costs at Region 5 in the imperfective sentences.

**Interaction Hypothesis.** Since comprehenders make more detailed simulations of events in their minds in the imperfective sentences versus the perfective sentences, only for the imperfective sentences will comprehenders’ reading times reflect the duration of an event described as short or long, while such difference of duration will not show up in the reading times for the perfective sentences. Accordingly, reading times will be longer at the region of the verb with grammatical aspect (Region 5) in the imperfective sentences, and this will be due to an interaction effect of grammatical aspect and event duration at Region 5.

**Focusing Hypothesis.** Since, in imperfective sentences, the processing costs caused by any of the mechanisms that underlie the other hypotheses relate to the ongoing nature of an event versus its completion, the processing costs will only be observed in Experiment 3 (Progressive), but not in Experiment 4 (Resultative). To restate this, the region for the verb with grammatical aspect (Region 5) will show longer reading times in the
imperfective sentences than in the perfective sentences only in Experiment 3, but not in Experiment 4.²⁶

5.2.6 Data analysis

5.2.6.1 Progressive (Experiment 3)

Selected Participants and Items. First, the participants’ accuracy rates on the confirmation questions were calculated based on the entire set of trials (including critical and filler items), with the idea that participants should be excluded if their accuracy rate was below 80%. No participants were removed due to inaccuracy (the average of the accuracy rates from 40 people was 96.6%). Second, the items’ accuracy rates on the confirmation questions were calculated based on responses from all participants, with the idea that critical items should be excluded if their accuracy rate was below 80%. No critical items were removed due to this criterion (the average accuracy on the 24 critical items was 97.9%). All remaining incorrect responses were filtered out for subsequent analyses.

Third, outlying reading times were removed; reading times that were over 2500 milliseconds were replaced with the value of 2500 milliseconds (32 reading times, i.e., 0.25 % of reading times were replaced) in the entire set of trials (including critical and filler sentences), with the idea that these times would skew the mean values and standard

²⁶ All test items were manipulated such that -teî- unambiguously had either a progressive or a resultative reading. Moreover, Japanese is a verb-final language, and reading times at the region of the verb with grammatical aspect (Region 5) are related to the global readings, should sentences be locally ambiguous. Hence, any differences of behaviors under this hypothesis will be observed in reading times at Region 5.
deviations for subsequent procedures. Two nil reading times (due to a possible failure of
the participants to press the spacebar) were removed. Fourth, outlying participants were
searched for by calculating the mean value and standard deviation of the reading times of
the entire set of trials (including critical and filler probes) for each participant, with the
idea that participants should be excluded if their mean reading time was above or below 3
standard deviations of the overall mean. No participants were removed due to this
criterion (the average of overall mean reading times from 40 people was 543
milliseconds). Fifth, outlying items were searched for by calculating the mean value and
standard deviation of reading times in the critical items from all participants, with the
idea that items should be excluded if their mean reading time was above or below 3
standard deviations of the overall mean. No items were removed due to this criterion (the
average of mean reading times for the 24 critical sentences was 552 milliseconds). As a
result, the data from 40 people (i.e., 10 participants for each of four conditions) were left
for subsequent analyses.

*Selected Response Times.* The mean value and standard deviation of the reading times
(including critical and filler items) were calculated for each participant. Trials were
replaced if their reading times were above or below 3 standard deviations from the value
for each participant. With this criterion, 1.98% of overall reading times on the entire set
of trials (including critical and filler probes) were replaced with either 2500 milliseconds
or 3 standard deviations above or below participants’ mean, whichever was applicable.
5.2.6.2 Resultative (Experiment 4)

**Selected Participants and Items.** First, the participants’ accuracy rates on the confirmation questions were calculated based on the entire set of trials (including critical and filler items), with the idea that participants should be excluded if their accuracy rate was below 80%. No participants were removed due to inaccuracy (the average of the accuracy rates from 20 people was 96.3%). Second, the items’ accuracy rates on the confirmation questions were calculated based on responses from all participants, with the idea that critical items should be excluded if their accuracy was below 80%. No critical items were removed due to this criterion (the average accuracy on the 24 critical items was 97.1%). All remaining incorrect responses were filtered out for subsequent analyses.

Third, outlying reading times beyond a fixed cutoff of 2500 milliseconds were removed, and such reading times were replaced with the value of 2500 milliseconds (5 reading times, i.e., 0.078% of reading times, were removed and replaced) on the entire set of trials, with the idea that outlying times would skew the mean values and standard deviations in subsequent procedures. Fourth, outlying participants were searched for by calculating the mean value and standard deviation of the reading times of the entire set of trials (including critical and filler probes) for each participant, with the idea that participants should be excluded if their mean reading time was above or below 3 standard deviations of the overall mean. No participants were removed due to this criterion (the average of mean reading times from 20 people was 530 milliseconds). Finally, outlying items were similarly searched for by calculating the mean value and standard deviation of reading times in critical items from all participants, with the idea that items should be excluded if their mean reading time was above or below 3 standard deviations of the
overall mean. No items were removed due to this criterion (the average of mean reading times on the 24 critical sentences was 536 milliseconds). As a result, data from 20 people (i.e., five participants for each of four conditions) were left for subsequent analyses.

**Selected Response Times.** The mean value and standard deviation of the reading times (including critical and filler items) were calculated for each participant. Trials were replaced if their reading times were above or below 3 standard deviations from the value for each participant. With this criterion, 1.94% of overall raw reading times on the entire set of trials (including critical and filler probes) were replaced with either 2500 milliseconds or 3 standard deviations above or below the participant’s mean, whichever was applicable.

5.2.6.3 Residual reading times (Experiments 3 and 4)

The process of converting the cleaned raw reading times to residual reading times was identical with the process used in Experiment 2 (English), with one exception regarding word order. Because the three critical regions in Experiments 3 and 4 were Regions 3, 4, and 5, a regression equation for each participant was constructed from word length using reading times from filler and experimental items by setting aside reading times at Regions 3, 4, and 5 (as was done for Regions 2, 3, and 4 in Experiment 2) of the critical items. Reading times used were those that were followed by correct responses to the confirmation questions. For filler items, two items with confirmation question response accuracy rates of 70% and 78% (i.e., below 80%) were removed in Experiment 3 (the average accuracy rate on the 48 filler items was 96%). No filler items were removed due
to this criterion in Experiment 4 (the average accuracy rate on the 48 filler items was 96%). In these Japanese experiments, en quad (used for Arabic numerals) and em quad (used for regular letters in Japanese including kanji characters and punctuation) were both counted as one letter. Residual reading times were obtained by subtracting each expected reading time, which was calculated by each regression equation, from the actual reading time for each participant.

In Experiment 3 (Progressive), the mean multiple R and the mean R square of regression equations across 40 people were 0.164, and 0.033, respectively. In Experiment 4 (Resultative), the mean multiple R and the mean R square of regression equations across 20 people were 0.138 and 0.0308, respectively. These values indicate that the strength of relation between word length and prenormalized reading times in Experiments 3 and 4, was, overall, 0.164 and 0.138, respectively, on the scale of 0 to 1 (with 1 the maximum), and that the equation can predict 3% perfectly the estimated reading times in both Experiments 3 and 4. These values mean that word length is not a good predictor of reading times in the Japanese experiments (Experiments 3 and 4). By comparison, the English experiment (Experiment 2) had 0.38 for the mean multiple R and 0.16 for the mean R square of regression.

5.2.7 Results

5.2.7.1 Progressive (Experiment 3)

Mean phrase-by-phrase raw reading times computed across participants are plotted in Figure 5.3 for each condition in each of the five regions. The presentation is simplified by
only graphing the raw RTs, as the critical tests showed no significant differences between raw and residual reading times (although the data patterns in the raw and residual times were not necessarily numerically identical). Table 5.5 presents a complete set of mean raw and residual reading times by conditions and by positions. Considering that the residual reading times are not based on a good estimate, I will discuss raw reading times as the main analyses, with residual reading times as sub-analyses.

Table 5.5. Raw (and residual) reading times per word in each region from Experiment 3, by participant, in milliseconds

<table>
<thead>
<tr>
<th>Condition</th>
<th>Region1</th>
<th>Region2</th>
<th>Region3</th>
<th>Region4</th>
<th>Region5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Imp- Long</td>
<td>521 (11)</td>
<td>572 (36.8)</td>
<td>582 (36.1)</td>
<td>508 (-19.6)</td>
<td>515 (-43.5)</td>
</tr>
<tr>
<td>Perf - Long</td>
<td>518 (8.2)</td>
<td>589 (54.9)</td>
<td>617 (69.4)</td>
<td>517 (-9.8)</td>
<td>507 (-21)</td>
</tr>
<tr>
<td>Imp- Short</td>
<td>526 (15.8)</td>
<td>567 (33.3)</td>
<td>596 (62)</td>
<td>504 (-22.9)</td>
<td>537 (-22)</td>
</tr>
<tr>
<td>Perf - Short</td>
<td>531 (20.7)</td>
<td>573 (38.2)</td>
<td>601 (67.8)</td>
<td>493 (-33)</td>
<td>516 (-13)</td>
</tr>
</tbody>
</table>

*Abbreviations: Imp = Imperfective, Perf = Perfective

Figure 5.3. Mean raw reading times in each region from Experiment 3, by participants, in milliseconds, and 95% confidence intervals for each condition group.
No differences were predicted for Region 1 (the subject) or for Region 2 (an adverbial), because they were not critical regions and all $F$s were less than 1, as expected. In the first critical region, Region 3 (the event duration phrase), there was no main effect of event duration (for both raw and residual reading times [RTs], both $F$s < 1 by participants and by items). There was no main effect of aspect (raw and residual RTs: both $F$s < 1.6 by participants and by items), or interaction between aspect and event duration (raw and residual RTs: both $F$s < 1.8 by participants and by items) in this region, as expected in a region prior to the region of verb and grammatical aspect.

In the second critical region, Region 4 (post event duration phrase for any spillover effects), the conditions again showed no difference for main effect of duration (raw and residual RTs: both $F$s < 1.5 by participants and by items). There was no main effect of aspect (raw and residual RTs: both $F$s < 1.2 by participants and by items), or interaction between aspect and event duration (raw and residual RTs: both $F$s < 1.2 by participants and by items), again as expected for a region prior to the region of verb and grammatical aspect.

In the last critical region, Region 5 (the verb with grammatical aspect), there was no main effect of aspect (raw and residual RTs: both $F$s < 1.8 by participants and by items), no main effect of duration (raw and residual RTs: both $F$s < 1.9 by participants and by items), and no interaction between aspect and event duration (raw and residual RTs: both $F$s < 1 by participants and by items); these results are counter to Uniform Hypothesis 1 and the Interaction Hypothesis, but not to Uniform Hypothesis 2.
5.2.7.2 Resultative (Experiment 4)

Mean phrase-by-phrase raw reading times computed across participants are plotted in Figure 5.4 for each condition in each of the five regions. Again, as in Experiment 3, the critical tests showed no significance between the raw and residual times, so the presentation is simplified by only graphing the raw reading times (although the data patterns in raw times were not necessarily numerically the same as those in residual times). Table 5.6 presents a complete set of mean raw and residual reading time by conditions and by positions.

![Figure 5.4. Mean raw reading times in each region from Experiment 4, by participants, in milliseconds, and 95% confidence intervals for each condition group](image)

<table>
<thead>
<tr>
<th>Condition</th>
<th>Region1</th>
<th>Region2</th>
<th>Region3</th>
<th>Region4</th>
<th>Region5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Imp - Long</td>
<td>529 (14.5)</td>
<td>546 (23.6)</td>
<td>514 (-12.5)</td>
<td>482 (-39)</td>
<td>527 (-22.7)</td>
</tr>
<tr>
<td>Perf - Long</td>
<td>536 (22.2)</td>
<td>564 (38)</td>
<td>549 (22)</td>
<td>496 (-24.3)</td>
<td>512 (-13.6)</td>
</tr>
<tr>
<td>Imp - Short</td>
<td>546 (31)</td>
<td>581 (57.5)</td>
<td>509 (-9.4)</td>
<td>507 (-14)</td>
<td>525 (-24.6)</td>
</tr>
<tr>
<td>Perf - Short</td>
<td>528 (14.9)</td>
<td>549 (25.8)</td>
<td>523 (3.9)</td>
<td>473 (-46.1)</td>
<td>528 (1.5)</td>
</tr>
</tbody>
</table>

Table 5.6. Raw (and residual) reading times per word in each region from Experiment 4, by participant, in milliseconds

*Abbreviations: Imp = Imperfective, Perf = Perfective
No effects were predicted for Region 1 (the subject) or Region 2 (an adverbial) as they are not critical regions, and as expected, the conditions did not differ at these regions (Region 1: all $F$s < 1; Region 2: all $F$s < 1.8). In the first critical region, Region 3 (the event duration phrase), there was no main effect of event duration (raw and residual RTs: both $F$s < 1.3 by participants and by items). There was no main effect of aspect (raw and residual RTs: both $F$s < 1.9 by participants and by items) or interaction between aspect and event duration (raw and residual RTs: both $F$s < 1 by participants and by items) in this region, as expected in a region prior to the region of verb and grammatical aspect.

In the second critical region, Region 4 (post event duration phrase for any spillover effects), the conditions again showed no difference in main effect of duration (raw and residual RTs: both $F$s < 1 by participants and by items), or main effect of aspect (raw and residual RTs: both $F$s < 1 by participants and by items). However, there was a suggestion of interaction, unexpected based on any of the hypotheses under discussion, between aspect and event duration in this region, but only by participants, with effects being marginal by items (raw RTs: $F_1 (1, 19) = 4.79, p = .041$, $F_2 (1, 23) = 1.76, p = .198$; residual RTs: $F_1 (1, 19) = 4.59, p = .045$, $F_2 (1, 23) = 3.188, p = .087$). This is likely a Type 1 error, which is to be expected given the number of tests I conducted. Paired $t$-tests further showed that the suggestion of interaction seems to come from the following: In short event conditions, the imperfective was read more slowly (raw mean times by participants = 506.59, $SD = 116.18$) than the perfective (raw mean times by participants = 472.78, $SD = 96.13$), but only by participants and not by items (raw RTs: $t_1 (19) = 2.33, p = .031$; $t_2 (23) = 1.40, p = .18$, two-tailed; residual RTs: $t_1 (19) = 2.19, p = 0.042$; $t_2 (23) = 1.47, p = .16$, two-tailed). The other paired $t$-tests did not show any
significance, with the following having the largest effects: In the imperfectives, there was 
a difference in that long event sentences were read faster (raw mean times by participants 
= 482.12, \( SD = 109.55 \)) than short event sentences (raw mean times by participants = 
506.59, \( SD = 116.18 \)), but the difference was without significance in either raw times or 
residual times and both by participants and by items (raw RTs: \( t_1(19) = -1.34, p = 0.20; \) 
\( t_2(23) = -0.90, p = 0.38 \), two-tailed; residual RTs: \( t_1(19) = -1.35, p = 0.19; t_2(23) = -1.52, \) 
\( p = 0.14 \), two-tailed).

In the last critical region, Region 5 (the verb with grammatical aspect), there was 
no main effect of aspect, or main effect of duration, or interaction between aspect and 
event duration (raw and residual RTs: all \( F_s < 1 \) by participants and by items); these 
results are counter to all of the hypotheses under consideration except for Uniform 
Hypothesis 2.

5.2.8. Discussion

5.2.8.1 Progressive (Experiment 3)

The results of the experiment reported here demonstrate that the imperfective sentences 
in Japanese, with verbs that evoke progressive readings, were read equally fast as the 
perfective sentences at every sentence position, with either long or short events.

Prior to Region 5 (the verb and grammatical aspect), null effects at Region 3 (the 
event duration) and Region 4 (the post event duration phrase for any spillover effects) 
conform to predictions based on the four hypotheses under discussion (i.e., Uniform 
Hypotheses 1 and 2, Interaction Hypothesis, and the Focusing Hypothesis), since any
effects of grammatical aspect, which may or may not interact with information on
duration of event, are expected to emerge as soon as readers encounter the grammatical
aspect at Region 5. It is worth noting that the null effects of event duration at Region 3
seem to conform to what was observed in Experiment 2 (English): Experiment 2 showed
a marginal effect of event duration at the equivalent region (Region 3, the event duration),
but the effect was local, without spillover effects, and only found in residual times,
suggesting the possibility that residual times had been overcalculated in the long-event-
duration conditions. The present results indicate that the readers of Japanese sentences
did not reflect differences between short or long event duration in their reading times at
the region for event duration.

In Region 5, there was no effect of grammatical aspect, which counters all of the
hypotheses under consideration except for Uniform Hypothesis 2. As briefly discussed in
regard to Experiment 2, the introduction of sentential objects that were not previously
mentioned or implied in the discourse may result in increasing reading times at the
position for the direct object, as the verb semantics need to integrate with the semantics
of the object (Coll-Florit and Gennari, 2011). As discussed in Section 5.1.8, the results in
Experiment 2 seem to be better interpreted as spillover effects from the preceding region.
However, it is now worth reconsidering Coll-Florit and Gennari’s explanation, because in
the Japanese experiments, the readers encounter the direct object prior to the verb and
grammatical aspect, and this raises the possibility of processing costs in those earlier
regions. A study by Frazier, Carminati, Cook, Majewski, and Rayner (2006) suggested
that the direct object noun may serve as a semantic confirmation of event interpretation as
completed or ongoing. Testing English garden path sentences, Frazier et al. found that
readers were more affected by the early closure penalty in reading the perfective (e.g., *While John hunted the frightened deer escaped*) than in reading the imperfective (*While John was hunting the frightened deer escaped*). Frazier et al. explained that the perfective grammatical aspect serves as semantic confirmation of the eventive/external perspective interpretation of the event (*to hunt*), which fits best with the presence of a definite description of an object (*the deer* versus *a deer or deers*). One must note that the word orders are different in Japanese and English, making any direct application of Frazier et al.’s idea to Japanese difficult. Furthermore, the critical sentences in Experiment 3 were not necessarily transitive sentences: Half of the 24 sentences were transitive (the DO versions), while the other half were mostly intransitive (the Adv [adverbial] versions).

Yet suppose that the presence of a direct object noun seen prior to the verb and grammatical aspect may serve as semantic confirmation for the perfective description of the event, in the reverse direction from what is expected in English. In the experimental sentences, the direct object noun in the DO version appeared at Region 3 with a definite event duration description (e.g., *12 plates for to count*). Readers may thus have expected the perfective more than the imperfective, possibly leading to longer reading times in the imperfective prior to Region 5 because of a mismatch with that expectation. The expected longer reading times in the imperfective sentences were as predicted by the other present hypotheses. However, no main effects of aspect were observed at any regions prior to Region 5, making this explanation unlikely.

Another explanation should also be considered for the other half of the critical items, the intransitive sentences (Adv version). Sentences in the Adv version were mostly intransitive; by analogy, in intransitive sentences the absence (vs. the presence) of the
A direct object might serve as a semantic confirmation of the ongoing nature of an event (vs. a completed event), which might fit with the imperfective sentences (vs. perfective sentences). However, the absence of a direct object may not fully confirm that an unfolding sentence is intransitive, because (a) the direct object is not obligatorily expressed in Japanese (i.e., object drop is an available option in Japanese), and (b) the option of scrambling allows any constituent to appear at almost any sentence position (i.e., possibly later in the sentence). In short, readers of intransitive sentences can expect the direct object at any sentence position unless they have already read the direct object. Therefore, the absence of the direct object may not necessarily lead readers to expect one grammatical aspect over the other prior to reading the grammatical aspect at Region 5. Rather, it is more likely that when the readers were presented with the verb and grammatical aspect at Region 5 they had no expectation of what they would read there. The lack of difference in reading times at any position before Region 5 does not support the possibility of a bias towards one grammatical aspect over the other prior to Region 5.

Yet another potential explanation for the null results may be that readers were not sensitive to the experiment. This possibility is, however, unlikely. Figure 5.5 (below) shows the raw reading times of the filler items, along with those from the critical items. The data pattern of the filler sentences differs from the pattern of the critical items. Paired $t$-tests on participant means and two-sample $t$-tests assuming unequal variances on items means at Region 3 revealed that filler sentences were read significantly faster than either of the four conditions by participants, and the difference was marginal by items.

---

27 It is worth recalling that Japanese sentences were presented phrase by phrase by using blank spaces rather than underbars, due to a restriction of the Japanese font system. This made readers blind to the sentence length of a given sentence.

28 $F$ tests showed $p < .05$, indicating the unequal variances between pairs.
(Imp-Long vs. Filler, $t_1(39)=2.72, p=.0097; t_2(33)=1.67, p=0.10$, two-tailed; Perf-Long vs. Filler, $t_1(39)=3.82, p=.00047, t_2(35)=2.95, p=0.0056$, two-tailed; Imp-Short vs. Filler, $t_1(39)=2.96, p=.0052, t_2 (32)=1.89, p=0.068$, two-tailed; Perf-Short vs. Filler, $t_1(39)=3.68, p=.000696, t_2 (32)=2.15, p=.039$, two-tailed). The difference in patterns of reading times between critical and filler items indicates that participants were sensitive to the experiment.

![Reading Times Graph](image_url)

Figure 5.5. Mean raw reading times in each region in the filler items from Experiment 3, by participants, in milliseconds, and 95% confidence intervals. Abbreviations: Imp = Imperfective, Perf = Perfective.

29 The comparisons by items were not direct, as there were 48 filler items and 24 critical items. Moreover, the filler items varied in length, ranging from 2 phrases to 6 phrases.
A final potential explanation for the null effects in Region 5 (and regions prior to Region 5) is the possibility that the presence of information on event duration may have removed processing costs that were expected in the imperfective conditions by informing readers of the boundary of the event so that they no longer wondered how long the event would continue. This fits Uniform Hypothesis 2, which predicts that the expected effects will emerge if the information on event duration is omitted. Of the four hypotheses, Uniform Hypothesis 2 is at this point most likely. This possibility is the central question that will be tested in Experiment 5.

5.2.8.2 Resultative (Experiment 4)
The only significant effect that was observed in the five regions was a suggestion of interaction between aspect and event duration at Region 4, but only by participants. This interaction further revealed that, in short event conditions, the imperfective was read more slowly than the perfective (but only by participants). Except for these suggestive effects, there were no other effects of aspect or event duration at any position, indicating that the Japanese imperfective sentences with verbs that evoke resultative readings (which is unique to Japanese) were read equally as fast as the perfective sentences, with either long or short events, giving no support for the Focusing Hypothesis.

As in Experiment 3 (Progressive), there were no effects of aspect or duration or interaction at Region 5 in Experiment 4, which is counter to Uniform Hypothesis 1 and the Interaction Hypothesis, but not to Uniform Hypothesis 2. Also similarly to Experiment 3, in Experiment 4 it may be a possibility that sentence constituents led readers to prefer one grammatical aspect over the other prior to reading the verb and grammatical aspect at Region 5. This possibility needs to be explored from a different
angle by discussing resultative sentences. As Section 3.3.2 discussed, resultative sentences in general take achievement verbs. The subject noun in such sentences often bears the theme role, which is related to the end-state of verbs. Inanimate entities are common as the subject noun phrase, along with animate entities for the theme, and the critical sentences used in the present study were not an exception: 11 out of the 24 critical items used inanimate entities as the subject. Inanimate entities are less agentive than animate entities. Therefore, an inanimate entity found at the subject position is likely to be taken as a theme. One possibility is that the theme role at the subject position may serve as a cue, similar to the direct object noun in the progressive sentences, because both are likely to bear a theme role, which is related to the end-state. The end-state implied by a theme role at the subject position may serve as semantic confirmation for the completed event situation, which then may fit with the perfective sentences. Crucially, though, the end-state is also compatible with the resultative use of the imperfective sentences, which places focus on the end-state. Hence, a potential preference on grammatical aspect by the readers may be one that is compatible with the end-state, which fits with either the perfective or (the resultative reading of) the imperfective. Therefore, only the grammatical aspect available at Region 5 can signal to the readers whether the situation (i.e., the resultant state of an event) is expressed as durational or punctual. The present results, however, showed no differences between the conditions at Region 5 or at any region prior to Region 5. Uniform Hypothesis 1 posits that imperfective sentences describe durative events while perfective sentences describe punctual events, and that durative events are costly as they elicit semantically more diverse associations, reflecting our experience, than do non-durative events. But the null results at Region 5 in the
present study seem to argue against this hypothesis, as the durativity of the described
situations of the resultative state in the imperfective did not result in any difference in
reading times from the perfective. Nor did reading times in the imperfective sentences
interact with a short or long description of the resultant state, giving no support for the
Interaction Hypothesis.

An explanation should also be considered for the other 13 critical items, which
used animate entities for the subject noun. Because the animate subject noun can be
agentive, and to the readers it may look no different from those used in progressive
sentences, an explanation for the results from such sentences may be similar to the
explanation discussed for the results from the progressive sentences in Experiment 3,
although the constituents might gradually signal that the unfolding sentence is more of a
resultative than a progressive sentence (e.g., by reason of probability and/or likeliness).
Yet the lack of differences between conditions at any region does not support this
possibility, either.

There was a suggestion of interaction observed at Region 4. However, it is very
unlikely that the interaction is suggestive of support for the Interaction Hypothesis. Even
if the discussed possibility of readers having a bias towards one grammatical aspect prior
to encountering the verb is the case, any such bias would be irrelevant as a cue on
whether to view a situation as durative or punctual, which is only signaled at Region 5.
Furthermore, the results showed no interaction at Region 3, hence the suggestion of
interaction at Region 4 seems not to be a spillover from Region 3, which is the earliest
region where the event duration was available to readers. Moreover, the effects did not
reach significance and were only marginal. It is therefore unlikely that grammatical
aspect interacts with event duration at Region 4, and hence these unexpected suggestive effects seem to be due to unknown factors irrelevant for the present hypotheses.

Supporting the sensitivity of the readers to the experiment, Figure 5.6 shows that the data patterns present differences between filler items and critical items in raw reading times (although they are subtle). For instance, the imperfective conditions were read more slowly than fillers at Region 5. Paired $t$-tests on participants’ means and two-sample $t$-tests on items means, assuming equal variances for the comparison of fillers in the
imperfective-short condition and unequal variances\(^{30}\) in the imperfective-long condition revealed significant differences (Imp-Long vs. Fillers: \(t_1\) (19) = 2.68, \(p = 0.015\), \(t_2\) (34) = 2.04, \(p = 0.049\), two-tailed; Imp-Short vs. Fillers: \(t_1\) (19) = 2.32, \(p = 0.032\), \(t_2\) (43) = 2.16, \(p = 0.036\), two-tailed).

The null results at Region 5 (and regions prior to Region 5) may therefore be taken as support for Uniform Hypothesis 2 as far as the present results go: It is possible that the readers did not show differences in reading times between the conditions because the information on event duration given prior to Region 5 may have eliminated the imperfective sentences’ expected cost to readers, who would otherwise be uncertain how long the resultant state continues. This possibility will be tested in Experiment 5.

5.3 Without event duration before grammatical aspect (Japanese)

Experiment 5 is identical to Experiments 3 and 4, with the exception that the region for event duration was removed, as a control experiment to Experiments 3 and 4.

5.3.1 Participants

Forty-six undergraduate students at Osaka University (in Toyonaka, Japan) participated in the experiment for compensation of $7 or $12. Participants were compensated $7 when they participated in this study alone, or $12 when they also participated in another experiment (Experiment 8 or 9) afterward. All participants were self-reported to be monolingual native speakers of Japanese.

---

\(^{30}\) F tests showed \(p > .05\), indicating the unequal variances between the pairs.
5.3.2 Materials

The materials used in Experiment 5 were identical to those used in Experiments 3 and 4, with two exceptions. First, the test items in Experiment 5 excluded the duration phrase from the original sentences. Second, half of the original items from each of Experiments 3 and 4 were used, rather than the full 24 items from each.

Critical Sentences. The 24 critical sentences were modified versions of 12 sentences used in the Progressive study (Experiment 3) and 12 used in the Resultative study (Experiment 4). All the sentences shared an identical structure with four phrases: the subject (Region 1), an adverbial (Region 2), another adverbial (Region 3), and a verb with grammatical aspect (Region 4). The critical region in this experiment was Region 4. Table 5.7 illustrates the structure of the critical sentences.

<table>
<thead>
<tr>
<th>Region of Interest</th>
<th>Region 1</th>
<th>Region 2</th>
<th>Region 3</th>
<th>Critical (V+Aspect)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regions</td>
<td>Atsushi-Top</td>
<td>yesterday</td>
<td>in his room</td>
<td>眠っていた。/眠った。 was sleeping./slept.</td>
</tr>
<tr>
<td>Sentence</td>
<td>淳は</td>
<td>昨日</td>
<td>彼の部屋で</td>
<td>眠っていた。/眠った。 was sleeping./slept.</td>
</tr>
</tbody>
</table>

The 12 items from the Progressive study were all adverbial versions. Eliminating the region for event duration (Region 3 in Experiment 3) from the D.O. version would eliminate the sentence constituent (i.e., D.O.). Moreover, the D.O. itself may help delimit an event, even if only the numeral is removed. Hence, the sentences from the D.O. versions were dispreferred for the purpose of the study. The 12 items from the Resultative study were those that used typical verbs for resultative readings and that
maintained naturalness when the phrase for event duration was removed. Of the
sentences that each participant saw, half had verb phrases with imperfective aspect (15),
and half had perfective (16). See Appendix F for a complete list of critical sentences used
in Experiment 5. The following example sentences illustrate the two conditions in an
adverbial version.

(15) Condition 1. Imperfective: 淳は/ 昨日/ 彼の部屋で/ 眠っていた。/

Gloss: Atsushi-Top/ yesterday/ in his room/ {sleep-tei-ta/ sleep-ta}/.
English Translation: ‘Atsushi was sleeping/slept in his room yesterday.’

**Filler Sentences.** The filler items used in Experiment 5 were identical to those used in
Experiments 3 and 4.

**Confirmation Questions.** The materials used in Experiment 5 were identical to those
used in Experiments 3 and 4.

**5.3.3 Procedure**
The procedure in Experiment 5 was identical to that in Experiments 3 and 4.

**5.3.4 Conditions**
Experiment 5 tested two conditions with one variable: grammatical aspect with two levels
(Imperfective, Perfective). Reading times at the sentence position of Region 4 (verb plus
grammatical aspect) were the dependent measures in this experiment. Reading times will
be analyzed separately for the Progressive items and the Resultative items.
5.3.5 Predictions

Experiment 5 tested one particular hypothesis, Uniform Hypothesis 2 (Uncertainty Hypothesis), that the on-line sentence processing of events may, or may not, be impacted by an interaction of information about event endedness and grammatical aspect. The other three hypotheses are not discussed here either because they would give identical predictions as in Experiments 3 and 4 (Uniform Hypothesis 1 and Focusing Hypothesis), or because it is not applicable to test interaction effects (Interaction Hypothesis) now that the region for event duration is removed. Specific predictions follow.

Uniform Hypothesis 2 (Uncertainty Hypothesis). Since the imperfective sentences leave readers uncertain about how long the event will last, unlike the perfective sentences, and readers are not informed of the event duration in this experiment, the imperfective sentences will take longer for readers to process than the perfective. Reading times at the region for the verb with grammatical aspect (our Region 4) will be longer in the imperfective sentences compared to the perfective sentences.

5.3.6 Data analysis

Selected Participants and Items. First, the participants’ accuracy rates on the confirmation questions were calculated based on the entire set of trials (including critical and filler items) with the idea that participants should be excluded if their accuracy rate was below 80%. No participants were removed due to inaccuracy (the average of the accuracy rates from 46 people was 96.6%). Second, the items’ accuracy rates on the
confirmation questions were also calculated based on the responses from all participants, with the idea that critical items should be excluded if their accuracy rate was below 80%. Although one critical item showed an accuracy rate of 78%, it was kept to maintain statistical power because (a) 78% was an acceptable rate, and (b) the confirmation question may have been a misleading one.\textsuperscript{31} Hence, no critical items were removed due to this criterion (the average accuracy on the 24 critical items was 97.2%). All remaining incorrect responses were filtered out for subsequent analyses.

Third, outlying reading times were removed at the fixed cutoff time of 2500 milliseconds, and such reading times were replaced with the value of 2500 milliseconds (36 reading times, i.e., 0.26% of reading times were removed and replaced due to this threshold) on the entire set of trials, with the idea that outlying times would skew the mean values and standard deviations in subsequent procedures. Fourth, outlying participants were searched for by calculating the mean value and standard deviation of the reading times of the entire set of trials (including critical and filler probes) for each participant, with the idea that participants should be excluded if their mean reading time was above or below 3 standard deviations of the overall mean. No participants were removed due to this criterion (the average of the mean reading times from 46 people was 594 milliseconds). Fifth, outlying items were searched for by calculating the mean value and standard deviation of reading times in the critical items from all participants, with the idea that items should be excluded if their mean reading time was above or below 3 standard deviations of the overall mean. No items were removed due to this criterion (the

\textsuperscript{31} The question was 哲也は静かでしたか？ ‘Was Tetsuya quiet?’ on the test item 哲也は窓の近くで静かにすすんていただけですか？ ‘Tetsuya was cooling/cooled off by the window in a quiet manner.’
average of mean reading times on the 24 critical sentences was 598 milliseconds). Finally, in order to have an equal number of participants for each condition, two participants (those with the lowest accuracy rates on the confirmation questions) were removed. As a result, data from 44 people (i.e., 11 participants for each of four conditions, because the two item sets were Latin squared) were left for subsequent analyses.

**Selected Response Times.** The mean value and standard deviation of the reading times (including critical and filler items) were calculated for each participant. Trials were replaced if their reading times were above or below 3 standard deviations from the value for each participant. With this criterion, 2.06% of the overall raw reading times on the entire set of trials (including critical and filler probes) were replaced with either 2500 milliseconds or 3 standard deviations above or below the participant’s mean, whichever was applicable.

**Residual Reading Times.** The process of converting the cleaned raw reading times to residual reading times was identical to that used in the other Japanese experiments, Experiments 3 and 4. Because there is only one critical region in Experiment 5 (i.e., Region 4), a regression equation for each participant was constructed from word length using reading times from filler and experimental items by setting aside reading times at Region 4 from the critical items. The reading times used were those that were followed by correct responses to the confirmation questions. Residual reading times were obtained by subtracting the expected reading times, calculated by regression equations, from the actual reading times by each participant.
The mean multiple R and the mean R square of regression equations across 44 people were 0.123, and 0.021, respectively. These values indicate that the strength of relation between word length and prenormalized reading times was 0.123 on the scale of 0 to 1 (with 1 the maximum), and that the equations can predict 2% of the estimated reading times. These values suggest that word length is not a good predictor of reading times in Experiment 5, as in the other Japanese experiments (Experiments 3 and 4).

5.3.7 Results

Mean phrase-by-phrase raw reading times computed across participants are plotted in Figure 5.7 for each condition in each of the four regions. The presentation is simplified by only graphing the raw reading times, as the critical tests showed identical patterns between raw and residual reading times (although the data patterns in the raw and residual times were not necessarily numerically identical). Table 5.8 presents a complete set of mean raw and residual reading times by condition and position. Considering that the residual reading times were not a good estimate, I will discuss raw reading times as main analyses with residual reading times as sub-analyses in what follows.

Separate paired t-tests were conducted on the set of the 12 progressive items and the set of the 12 resultative items from the 44 participants. No tests are reported for Region 1 (the subject), Region 2 (an adverbial), or Region 3 (an adverbial) from either the progressive or resultative item sets, as these were not critical regions and the conditions did not differ at these regions in either raw or residual reading times.
Figure 5.7. Mean raw reading times in each region from Experiment 5, by participants, in milliseconds, and 95% confidence intervals for each condition group. Abbreviations: Prog = Progressive, Resul = Resultative, Perf = Perfective, Imp = Imperfective.

Table 5.8. Raw (and residual) reading times per word in each region from Experiment 5, by participant, in milliseconds

<table>
<thead>
<tr>
<th>Condition</th>
<th>Region1</th>
<th>Region2</th>
<th>Region3</th>
<th>Region4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prog-Perf</td>
<td>620 (52.7)</td>
<td>611 (23.4)</td>
<td>562 (-22)</td>
<td>558 (-29.4)</td>
</tr>
<tr>
<td>Prog-Imp</td>
<td>627 (59.9)</td>
<td>605 (16.3)</td>
<td>584 (0.2)</td>
<td>568 (-45.3)</td>
</tr>
<tr>
<td>Resul-Perf</td>
<td>657 (75.2)</td>
<td>626 (42.9)</td>
<td>533 (-52.4)</td>
<td>518 (-68.5)</td>
</tr>
<tr>
<td>Resul-Imp</td>
<td>663 (81.3)</td>
<td>638 (54.3)</td>
<td>536 (-49.2)</td>
<td>547 (-66.2)</td>
</tr>
</tbody>
</table>

*Abbreviations: Prog = Progressive, Resul = Resultative

*Progressive.* In Region 4 (the verb with grammatical aspect), comparison of raw reading times in the imperfective (mean by participants = 568.44, $SD = 145.79$) and the perfective (mean by participants = 558.38, $SD = 136.83$) revealed no significant differences between the conditions by participants or by items ($t_1(43) = -0.55$, $p = 0.585$, $t_2(11) = -0.81$, $p = 0.435$, two-tailed). The null differences were confirmed in residual reading times in the
imperfective (mean by participants = -45.33, SD = 102.70) and the perfective (mean by participants = -29.38, SD = 121.03) both by participants and by items ($t_1$ (43) = 0.86, $p$ = 0.397, $t_2(11) = 1.13, p = 0.281$, two-tailed).

**Resultative.** In Region 4 (the verb with grammatical aspect), comparison of raw reading times in the imperfective (mean by participants = 546.96, $SD = 174.98$) and the perfective (mean by participants = 518.07, $SD = 122.32$) revealed no significant differences between the conditions by participants or by items ($t_1$ (43) = -1.32, $p$ = 0.195, $t_2(11) = -1.28, p = 0.227$, two-tailed). The null differences were confirmed in residual reading times in the imperfective (mean by participants = -66.16, $SD = 121.89$) and the perfective (mean by participants = -68.53, $SD = 108.96$) both by participants and by items ($t_1$ (43) = -0.12, $p$ = 0.908, $t_2(11) = -0.20, p = 0.842$, two-tailed).

5.3.8 Discussion

In Experiment 5, the phrases that provided information on event duration in Experiments 3 and 4 were removed from the experimental sentences, and thus comprehenders were left uncertain about event duration. Nevertheless, the results of Experiment 5 showed that there was no increase in reading times in the imperfective condition versus the perfective condition at any position. Uniform Hypothesis 2 predicts that in the absence of information on event duration, reading times on imperfective sentences should be longer than on perfective sentences at the critical region. The results of Experiment 5, however, revealed no differences between the conditions, and thus do not support Uniform Hypothesis 2. Nor do the results support either the Focusing Hypothesis (as there were no
differences due to focusing in event parts manipulated by progressive or resultative item sets) or Uniform Hypothesis 1, which predicted that the imperfective would be read more slowly than the perfective, regardless of the presence or absence of the event duration information.

In Region 4 (the region for verb with grammatical aspect), there were no effects of aspect. One potential explanation may be that, as discussed in regard to the other Japanese experiments, information that unfolded prior to Region 4 may have led readers to prefer one grammatical aspect over the other before actually encountering the grammatical aspect at Region 4. In this case, differences between the conditions would have appeared earlier. This possibility is again unlikely, because there were no significant differences between the conditions at any position prior to Region 4, in either progressive or resultative items. Compared to Experiments 3 and 4, the information presented prior to the verb and grammatical aspect was reduced by one phrase, so that any effects of information available prior to the verb and the grammatical aspect may have been reduced; however, the removal of the phrase does not change the explanation if it reduces any influence by reducing the amount of information available.

It is an unlikely possibility that readers were not sensitive to the experiment. Figure 5.8 shows the raw reading times of the filler items along with those from the critical items. Figure 5.9 additionally presents clearer data patterns that can be seen in the residual reading times. The data pattern of filler sentences in Experiment 5 differs from that of the critical resultative items, but not appreciably from that of the critical progressive items; the data pattern of the progressive sentences differs from that of the resultative items, indicating that participants were sensitive to differences between the
item sets. For instance, 2 x 2 two-factor ANOVAs with repeated measures on one factor (by items), and on both factors (by participants), revealed that the progressive items were read faster than the resultantive items at Region 3 (raw reading times [RTs]: $F_1(1,43) = 6.40, p = 0.015, F_2(1, 22) = 5.12, p = 0.034$; residual RTs: $F_1(1,43) = 6.94, p = 0.012, F_2(1,22) = 7.33, p = 0.013$) and marginally faster at Region 4 (raw RTs: $F_1(1,43) = 5.07, p=0.0296, F_2(1,22) =3.49, p=0.075$; residual RTs: $F_1(1,43) =4.77, p=0.035, F_2(1,22) = 3.55, p=0.073$). An ANOVA additionally showed a suggestion of main effect of aspect at Region 4 (the verb with grammatical aspect) but only in raw reading times (raw RTs: $F_1(1, 43) = 3.077, p = 0.087, F_2(1, 22) = 2.28, p = 0.145$; residual RTs: both $F$s < 1), indicating that the imperfective sentences were read more slowly than the perfective sentences, although the difference was without significance. Considering that the effect disappeared in the residual times, longer reading times in the imperfective condition were likely a reflection of the word length, which was two letters longer in the imperfective due to the presence of the morpheme -tei- てい (e.g., in 眠っていた ‘was sleeping’), which was absent in the perfective (e.g., 眠った ‘slept’). Other regions (non-critical) did not show any difference across the conditions, or showed a suggestion only in the raw times (not in the residual times), or only by participants (not by items).

Overall, the results of Experiment 5, which tested Japanese sentences, did not support any of the present hypotheses.

---

32 These factors were used because the progressive vs. resultative distinction was within participants and between items, and the imperfective vs. perfective distinction was within participants and within items (recall that Experiment 5’s 24 items consisted of 12 progressive and 12 resultantive items).
Figure 5.8. Mean raw reading times in each region in the filler items from Experiment 5, by participants, in milliseconds, and 95% confidence intervals. Abbreviations: Prog = Progressive, Resul = Resultative, Perf = Perfective, Imp = Imperfective.

<table>
<thead>
<tr>
<th>Region1</th>
<th>Region2</th>
<th>Region3</th>
<th>Region4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prog-Perf</td>
<td>620</td>
<td>611</td>
<td>562</td>
</tr>
<tr>
<td>Prog-Imp</td>
<td>627</td>
<td>605</td>
<td>584</td>
</tr>
<tr>
<td>Resul-Perf</td>
<td>657</td>
<td>626</td>
<td>533</td>
</tr>
<tr>
<td>Resul-Imp</td>
<td>663</td>
<td>638</td>
<td>536</td>
</tr>
<tr>
<td>Filler</td>
<td>642</td>
<td>611</td>
<td>573</td>
</tr>
</tbody>
</table>

Figure 5.9. Mean residual reading times in each region in the filler items from Experiment 5, by participants, in milliseconds, and 95% confidence intervals. Abbreviations: Prog = Progressive, Resul = Resultative, Perf = Perfective, Imp = Imperfective.

<table>
<thead>
<tr>
<th>Region1</th>
<th>Region2</th>
<th>Region3</th>
<th>Region4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prog-Perf</td>
<td>52.7</td>
<td>23.4</td>
<td>-22.0</td>
</tr>
<tr>
<td>Prog-Imp</td>
<td>59.9</td>
<td>16.3</td>
<td>0.2</td>
</tr>
<tr>
<td>Resul-Perf</td>
<td>75.2</td>
<td>42.9</td>
<td>-52.4</td>
</tr>
<tr>
<td>Resul-Imp</td>
<td>81.3</td>
<td>54.3</td>
<td>-49.2</td>
</tr>
<tr>
<td>Filler</td>
<td>56</td>
<td>21</td>
<td>-26</td>
</tr>
</tbody>
</table>
5.4 General discussion

The experiments reported in this chapter demonstrate non-parallel results for English and Japanese. In English, the imperfective sentences were read more slowly than the perfective sentences at the position where readers read the verb with the grammatical aspect and at the successive region, irrespective of whether or not the described event had a short or long event duration (Experiment 2). This study replicates an experiment in Spanish by Coll-Florit and Gennari (2011), and to my knowledge, is the first study to conduct such an experiment in English. In Japanese, however, the imperfective sentences were read as slowly as the perfective sentences at every position, irrespective of whether or not the described event had a short or long event duration. Effects were null when the information on event duration was given prior to the grammatical aspect (Experiments 3 and 4), and when that information was not given at all (Experiment 5). Effects were null irrespective of whether or not the described event had a progressive reading (Experiments 3 and 5) or a resultative reading (Experiments 4 and 5). Overall, these results indicate that (a) the difference in event description expressed by manipulating grammatical aspect has immediate consequences in on-line language comprehension in English, and (b) such processing activity, for some reason, does not take place in Japanese, at least on the surface level; Japanese readers do not appear to change their processing speed in reading sentences with either grammatical aspect.

The series of experiments was designed to test a set of hypotheses based on the assumption that the effects of aspect would be present in any language, but the empirical evidence showed that while such effects were present in English, they were absent in
Japanese. Because the study confirms the effect of grammatical aspect in English, neither the effects nor the hypotheses can be denied. Rather, the null effects in the Japanese experiments must be taken as an indication that the role of grammatical aspect in sentence processing depends on the language.

One question raised by these experimental results is why durative events took longer to process than non-durative events in on-line comprehension in English, as well as in Spanish according to Coll-Florit and Gennari (2011), but not in Japanese. What processing mechanism would account for such an effect in languages such as English and Spanish, but not in a language such as Japanese?

One potential explanation of the difference is that the relative timing of reading the grammatical aspect in a given sentence may have a direct impact on the processing cost of constructing situation models. In English, the grammatical aspect is processed at an earlier stage in constructing situation models. Readers have processed only the subject noun before they process grammatical aspect, which is marked on the main verb, before the direct object, which is claimed to be responsible for event telicity (Borer, 2005; Levin & Rappaport-Hovav, 1995 among many others. Also see Folli & Harley, 2006). In addition, any be verb that precedes the main verb in imperfective sentences is another cue biased toward the imperfective over the perfective (although this consideration should have minimally affected the present phrase-by-phrase reading task, where readers saw be-V-ing or V-ed in the conditions of imperfective vs. perfective). In contrast, on the assumption that sentence processing is incremental (e.g., Altmann & Steedman, 1988; MacDonald, Pearlmutter, & Seidenberg, 1994), Japanese readers would have created identical situation models in both conditions before they processed the grammatical
aspect, which does not appear until the sentence-final position. Thus, in creating incremental situation models English speakers expect more cues to come at the point they read grammatical aspect, while for Japanese speakers, when they read the grammatical aspect with the verb, it is a cue that no more information is to come. Therefore, one possibility is that part of the cost caused by grammatical aspect in a language like English is that comprehenders must remain prepared to integrate more information into the incremental situation model, while this cost would not be present in a language like Japanese.

Another potential explanation for how grammatical aspect in a given sentence may have an impact on the processing cost of constructing situation models is in terms of accessing knowledge related to the situation. The linguistic theories that Coll-Florit and Gennari (2011) discuss to account for the observed effects in their study all propose that situations are understood by activating stored experiential knowledge associated with the situations in question. One theory accounts for the effect by suggesting that durative states are inherently more difficult to process simply because they are longer and thus take longer to comprehend, by mentally re-enacting the events described (Barsalou, Simmons, Barbey, & Wilson, 2003; Zwaan, 2008; Zwaan & Taylor, 2006). Another theory suggests that “durative states take longer to process because their associated meanings engage a variety of features and related knowledge, ultimately derived from experience of the world and the language” (Coll-Florit & Gennari, 2011, p. 67). The former proposal is in line with simulation theories of processing, according to which comprehenders mentally re-enact the events described. In this approach, representations retain an analogical relation to multi-modal experiences of events (Matlock, 2004;
Matlock, Ramscar, & Boroditsky, 2005; Richardson & Matlock, 2007; Zwaan, 2008). The latter theory is grounded in attested probabilistic processing mechanisms and the claim is that the diverse/distributed nature of associated knowledge can give rise to processing mechanisms: “since durative expressions have diverse associations and occur in diverse contexts, any given association is likely as strong as any other, giving rise to difficulty in comparison with the strong associations of punctual expressions” (Coll-Florit & Gennari, 2011, p. 67). Coll-Florit and Gennari state that both theories can be maintained, and that more research is needed to conclusively support either alternative.

According to either theory, the timing of the readers’ encounter with grammatical aspect would influence processing costs differently between the two grammatical aspects, which allow stored experiential knowledge associated with the events to be activated to different degrees. This point is crucial for understanding the observed differences in processing Japanese, in which the grammatical cue is provided sentence-finally, and in processing English and Spanish, in which the grammatical cue is provided earlier in the sentence. Comprehenders re-enact the described events online as they continue to read the sentence: The more time comprehenders have while processing an unfolding sentence after they encounter grammatical aspect, the more associations with the situation may arise. Thus, in Japanese, the reason that processing costs do not differ between the grammatical aspects may be because comprehenders do not receive the cue that grammatical aspect provides about how to see the situation when their situation models are still mostly yet to be built.

Therefore, the grammatical aspect itself may not be costly. Rather, what makes sentences with different grammatical aspects have more or less processing cost may be
the role of grammatical aspect in providing a cue on how to structure situation models by having readers more or less prepared for incoming information, and/or influencing the process of how much readers associate the given situation with other situations.

In the Japanese experiments, a change in focus on certain parts of the events (an ongoing part vs. the end-state) did not manifest effects (Experiments 3, 4, and 5). Experiments 3 and 4, which tested the Focusing Hypothesis, and Experiment 5, which tested Uniform Hypothesis 2, may have had null results because the sentence-final cue for what viewpoint from which to see the described event may not greatly influence the situation-specific construction of situation models up to the final region. The null effects provide no clues as to whether the difference in focus in event parts or the information on event duration provided before the grammatical aspect may have had any effect. One might well expect that sentence-final aspect could have an effect on subsequent processing, such as for the next sentence in a discourse. This is not to say that the aspect in a sentence could have an effect on the construction of representations of an individual event(s) in the next sentence. Rather, readers construct a mental model of a clause, and then monitor whether there are any updates to make on it as they read the next clause. As Zwaan, Langston, and Graeser (1995) illustrate, “for example, if a clause indicates a time shift compared with the previous clause, then the temporal index of the model needs to be updated.” My assumption is that Japanese comprehenders would have completed the construction of the model of a described event by the time they read the information at the sentence-final position, and it is within this construction process that the aspect information could have any effect. Such a current event representation would then be integrated into the unfolding mental representation in subsequent processing, with
representations of individual events and actions being the building blocks of situation models (see Zwaan, 2008; Zwaan, Langston, & Graesser, 1995; Zwaan, Magliano, & Graesser, 1995; Zwaan & Radvansky, 1998; see also Speer, Zacks, & Reynolds, 2007; Speer & Zacks, 2005).

Further research possibilities include manipulating the position of grammatical aspect in Japanese sentences. For example, grammatical aspect can appear earlier in sentences with relative clauses like 眠っていた猫が (sleep-tei-Pst-cat-Nom ‘a cat that was sleeping’) or 眠った猫が (sleep-Pst-cat-Nom ‘a cat that slept’), so such sentences would allow the testing of aspect effects in Japanese. Transitive verbs can also appear in relative clauses such as 太郎が飼っていた猫が (Taro-nom keep-tei-Pst-cat-Nom ‘a cat that Taro was keeping’) or 太郎が飼った猫が (Taro-nom keep-Pst-cat-Nom ‘a cat that Taro kept’). Here, it is Taro that engages in the action, not the cat. In these transitive versions, the direct object will be processed after the grammatical aspect, as in sentences in English. If grammatical aspect plays a role by providing a cue on how to structure situation models by having readers more or less prepared for incoming information, then the processing cost in the imperfective can be expected to be larger than that of the perfective in such sentences. An additional thought about these possible test examples is that there is a shift of protagonist in the transitive version (to Taro, who engages in the action in the relative clause, from the cat, who is the protagonist of the main clause indicated by a nominative case -ga), while there is no such shift in the intransitive version (the cat is the protagonist in both the main and relative clauses). If simulations can be understood as activating stored experiential knowledge associated with the events in question (Coll-Florit and Gennari, 2011), the identical protagonist in the intransitive
version may allow more continuity in the knowledge activated by the clauses, as opposed
to the transitive version with the shift in protagonist. This matter should be discussed
with a wider focus than the current study’s focus on grammatical aspect and its role in
processing, and so I will not discuss it more here.

For Uniform Hypothesis 2, the null results of the experiments reported in this
chapter, especially Experiments 3 through 5, make any conclusion or interpretation
difficult. Still, Experiment 2 (English) provided a suggestive piece of evidence against
Uniform Hypothesis 2: It showed no decrease in reading latency in the imperfective
condition at the position where the reader read the event duration. It was predicted that, if
the hypothesis were true, once the reader was no longer uncertain about how long the
described event continued, the reading times would be reduced in the imperfective
compared to the perfective, but no such interaction was observed.

The Interaction Hypothesis was not supported. Given the theory that
comprehenders make more detailed simulations of events in their minds when the event
is described with the imperfective as opposed to the perfective (e.g., Bergen & Wheeler,
2010; Anderson et al., 2010), it was predicted that only for the imperfective sentences
would comprehenders’ reading times reflect the duration of an event described as short or
long, while such difference of duration would not show up with the perfective sentences.
However, although the results showed some suggestion that long event phrases took
longer to read than short event phrases (Experiment 2), information on event duration did
not show any interaction with aspect. Hence, the duration of event did not appear to be
reflected in situation-specific comprehension in either the imperfective or the perfective.
These results seem to counter approaches that suggest time is an abstract domain that must be understood with reference to concrete domains (Boroditsky, 2000; Boroditsky & Ramscar, 2002; Matlock et al., 2005). A possible explanation may be that information on event duration is less relevant to sensory-motor properties and thus fails to emerge in the present study, in contrast to previous studies. The previous works that have reported an interaction of event description and aspect were more related to the notion of sensory-motor action. For instance, Action-sentence Compatibility Effect (ACE) methodology (Glenberg & Kaschak, 2002) was explored by Bergen and Wheeler (2010), who manipulated the grammatical aspect of sentences about hand motion. They showed that progressive sentences facilitate manual action in the same direction, while perfect sentences do not. Another study, by Anderson, Matlock, and Spivey (2010), manipulated a description of terrain (difficult or easy) and the grammatical aspect of sentences about motion events. This study revealed that when comprehenders moved a character (by moving a computer mouse) on a terrain depicted on a computer screen in order to show a motion they heard described, the action took longer in the difficult terrain conditions when the sentence was expressed in the imperfective. In these two studies, sentences described actions related to sensory-motor properties, and in both studies the descriptions were responded to by a motion by the participants in experiments. In the current study, short or long event duration was cued by numerals, and was responded to by hitting a key. Numerals may be too abstract to interact with sensory-motor properties (compared to direction of motion or an description of ease or difficulty of paths for motion events). Moreover, our participants did not have to make motions similar to described motions. Hence, further studies might explore whether a notion of event
duration expressed by an abstract notion such as numerals may be reflected in sentence processing if a sensory-motor action is required to respond, to test the hypothesis that time is an abstract domain that must be understood with reference to concrete ones (Boroditsky, 2000; Boroditsky & Ramscar, 2002; Matlock et al., 2005).

As no main effect of event duration was confirmed in either the English or the Japanese experiments, this study takes a neutral position on whether any effects of event duration are dissociated from causality. Previous studies have found that causally complex events take longer to process than causally simpler ones (Gennari & Poeppel, 2003; McKoon & Macfarland, 2002) by manipulating causal structure. Following Coll-Florit and Gennari’s (2011) suggestion, the current experiment distinguished causality from duration by keeping causal structure constant while varying duration, namely by manipulating the duration of identical events; previous studies manipulated lexical items (i.e., verbs) to manipulate duration of events (e.g., Magliano & Schleich, 2000). In our study, long events were read more slowly than short events in Experiment 2 (English), but this main effect of event duration was local (it was limited to the region that presented the event duration phrase) and the effect was seen in residual times, but not in raw times, making it plausible that the effect is more an artifact of the residual-time calculation. Furthermore, no interaction effect was seen in the region of the event duration phrase, and thus the increase in reading times in long duration events was isolated from any effect of aspect. Overall, we have no consistent evidence of longer reading times in the long event condition than in the short event condition. Further studies are needed to address the question of whether the null effects in the current study are due to the use of the abstract notion of event duration, as discussed above.
In sum, this study’s results supported Uniform Hypothesis 1, which is the hypothesis equivalent to that proposed by Coll-Florit and Gennari (2011). Imperfective grammatical aspect was observed to be more costly than perfective aspect in the English experiment (Experiment 2) but not in the Japanese experiments (Experiments 3, 4, and 5). As discussed in this section, these results may be best interpreted if we consider the relative timing of readers’ processing of grammatical aspect, whether in terms of incremental processing that expects more or less upcoming information to integrate, or in terms of more or less elicitation of associations with the situation grounded in linguistic and world knowledge. By comparing the results of English and Japanese experiments, this study overall suggests that it is not grammatical aspect itself that is costly, but rather the role it plays in the process of constructing situation models and associating event knowledge during on-line sentence comprehension. Questions remain for future studies. For instance, a study should address the question of whether highlighting event parts (an ongoing state vs. the end-state) relates to any costs of reading imperfective sentences by positioning the grammatical aspect at an earlier stage in sentential processing (by use of the relative clause, as suggested in this section). Another study should address how we engage in comprehension of the abstract notion of time, and how that might interact with the way we process events in sentence comprehension.
Empirical studies have indicated that grammatical aspect plays an important role in memory retrieval. Carreiras, Carriedo, Alonso, and Fernández (1997) and Magliano and Schleich (2000) demonstrated that the imperfective makes a sentence’s information more accessible for memory retrieval compared with the perfective aspect (see Section 3.1.1). However, as is the case with many psycholinguistic studies on grammatical aspect, an uninvestigated question is whether the observed effects are from grammatical aspect or lexical aspect, or even from their interaction. One approach to this question is to ask whether highlighting on parts of an event relates to the role of the imperfective in memory retrieval (Focusing Hypothesis). As an instance of interaction between lexical and grammatical aspect, Section 3.3.2 discussed how the imperfective aspect in Japanese, -tei-, can uniquely highlight different parts of a described event as an outcome of its interaction with lexical aspect: Unlike in English, the imperfective evokes resultative meaning when it is used with achievement verbs, while it evokes progressive meaning with other verbs such as activity verbs and accomplishment verbs (Shirai, 2000). In the progressive reading, the imperfective highlights the ongoing nature of an event and provides temporal information about the unfolding of events, while in the resultative reading it highlights the completed, rather than ongoing, nature of an event.

Crucially, Ferretti et al. (2007, 2009) suggest that highlighted event parts are related to comprehenders’ retrieval of information (see Section 3.1.1). In Japanese, the
single form of the imperfective -tei- can correspond to two distinct focuses on event parts. Hence, by testing grammatical aspect in Japanese, it is possible to explore whether grammatical aspect plays its role in relation to what part of an event is focused on. The following subsection summarizes a hypothesis with which to approach the present question.

**Focusing Hypothesis.** The imperfective expresses the ongoing nature of a situation, rather than its completion, and the role of the imperfective in memory retrieval for event representation is to keep an event at a relatively high state of activation during subsequently evoked context (Magliano & Schleich, 2000). For this reason, the effects of the imperfective are observed when imperfective sentences highlight the ongoing nature of an event, but they are not observed, or they are short-lived, when they highlight an event’s completion. To restate this hypothesis in terms of the case of Japanese imperfective -tei-, the effects are observed only when the imperfective sentences express the progressive readings, but they are either not observed or they are short-lived when the imperfective sentences express the resultative readings, since the imperfective sentences that evoke resultative readings highlight the completed, and not the ongoing, nature of an event.

This chapter describes four experiments (Experiments 6 through 9) that explore whether the imperfective grammatical aspect helps memory retrieval of an event representation in comparison to the perfective grammatical aspect in on-line sentence processing of Japanese sentences. In each of the four experiments, participants read 24
short stories that included a critical sentence in either the imperfective (-tei-) or the perfective (-ta) grammatical aspect. Later in the stories, the readers encountered a critical verb as a probe word and answered the question of whether it had appeared in the preceding sentences in that story. The grammatical aspect was manipulated in all four experiments. The relative timing of memory retrieval of an event was manipulated by probe positions, by using probes closer to the critical sentence (Experiments 6 and 7), or by using probes further from the critical sentence (Experiments 8 and 9). The classes of verbs were manipulated between experiments, with verbs that evoke progressive readings (Experiments 6 and 8) and with those that evoke resultative readings (Experiments 7 and 9) to explore the role of difference in focus on event parts in sentence processing.

Considering the nature of memory decay, a faster reaction to probe recognition is predicted at a near probe position than at a far probe position across all experiments. Imperfective sentences help readers with memory retrieval, according to Magliano and Schleich (2000), but as this may or may not depend on what part of an event is highlighted, I have made separate predictions for each experiment.

Section 6.1 presents Experiments 6 (Progressive) and 7 (Resultative), which explored memory retrieval in a short span. Since the progressive use of the imperfective is equivalent to the English imperfective tested by Magliano and Schleich, I predict that Experiment 6 (Progressive) will replicate Magliano and Schleich’s (2000) results, because the imperfective helps memory retrieval of an event representation, in contrast to the perfective. Since the role of the imperfective is to maintain information in working memory, the perfective and imperfective conditions will have essentially the same level of activation at the near probe; but at the far probe, the imperfective will evoke faster
responses to the probe than the perfective. Experiments 6 and 7 test memory retrieval in a shorter span than those that were tested by Magliano and Schleich. Since Magliano and Schleich showed stronger effects in a longer span than in their shorter span conditions, I predict that the expected effects will be weaker in Experiment 6 (Progressive, short span) than in Experiment 8 (Progressive, long span). If the imperfective aspect uniformly plays a role in memory retrieval, i.e., regardless of what verb type it accompanies and which part of the event it highlights, the current prediction will also apply to Experiment 7 (Resultative) (*Uniform Hypothesis*).\(^{33}\)

However, if highlighting the ongoing nature of an event versus its completion plays a role in helping memory retrieval, effects of the imperfective in Experiment 7 (Resultative) will be non-existent or short-lived. In experiments that test memory retrieval in the short span (i.e., Experiments 6 and 7), the *Focusing Hypothesis* predicts that the results will not differ when the imperfective highlights an ongoing state and an end-state (Experiments 6 and 7, respectively), since the short-lived effects that are expected in Experiment 7 (Resultative) should survive for a short span in memory.

Section 6.2 presents Experiments 8 (Progressive) and 9 (Resultative), which explored memory retrieval in a long span. Since this long-span condition is identical to the condition tested by Magliano and Schleich, and since the progressive use of the imperfective is equivalent to the English imperfective tested by Magliano and Schleich, I predict that Experiment 8 (Progressive) will replicate the results of Magliano and

---

\(^{33}\) The *Uniform Hypothesis* that I discuss in Chapter 6 is essentially identical to *Uniform Hypothesis 1* that I discuss in Chapter 5, since they both assume that grammatical aspect plays its role uniformly with any event. They differ in their specifics, by considering effects for memory retrieval, or those for processing cost, respectively. The *Uniform Hypothesis 2* that I discuss in Chapter 5 is not relevant in Chapter 6, since it is a hypothesis on processing costs, and not on memory retrieval.
Schleich, i.e., showing that the imperfective helps memory retrieval of an event representation more than the perfective does, which will be demonstrated by faster responses at the far probe, since the role of the imperfective is to maintain information in working memory. If the imperfective aspect uniformly plays a role in memory retrieval, the current prediction also will apply to Experiment 9 (Resultative) (*Uniform Hypothesis*).

In contrast, if the highlighting of parts of events does relate to memory retrieval, predictions for Experiment 9 (Resultative) will diverge from those for Experiment 8 (Progressive). If the Focusing Hypothesis is correct, then I predict that, unlike in Experiment 8 (Progressive), in Experiment 9 (Resultative) there will be no effect or reduced effects of the imperfective. Namely, the Focusing Hypothesis predicts that the resultative use of the imperfective in the long span condition (Experiment 9) will show a reduced effect of helping memory retrieval compared to the short span condition (Experiment 7), on the assumption that effects from grammatical aspect with an end-state focus are short-lived.

### 6.1 Memory retrieval in short span

Experiments 6 and 7 were identical, with the exception that Experiment 6 focused on, and was conducted with, verbs that evoke progressive readings of the imperfective sentences while Experiment 7 focused on, and was conducted with, verbs that evoke resultative readings of the imperfective sentences.
6.1.1 Participants (Progressive, Resultative)

Eight undergraduate students at Tohoku University (in Sendai, Japan) participated in the Progressive study (Experiment 6), and a separate group of 40 undergraduate students at Tohoku University participated in the Resultative study (Experiment 7), for compensation of $7 or $12. Participants were compensated $7 when they participated in this study alone, or $12 when they participated in either one of the present studies after they had participated in another study (Experiments 3 or 4). All were self-reported to be monolingual native speakers of Japanese.

6.1.2 Materials (Progressive, Resultative)

Critical Stories. The 24 short stories were created by the researcher. Following Magliano and Schleich (2000), all the stories shared an identical structure that consisted of three components, i.e., introduction, experimental, and conclusion. The introduction component had three sentences that introduced the protagonist, and/or necessary background that preceded a situation described in the experimental component. The experimental component began with the critical sentence, which was followed by three post-critical sentences. A near or far critical probe was presented in this experimental component. The conclusion component contained two to three sentences to conclude the story.

Following Magliano and Schleich, the post-critical sentences were created such that they made sense with the critical sentence in either grammatical aspect (imperfective

\[34\] No heavy fatigue from participating in either of these experiments was assumed, since it was a relatively short five-minute experiment. Care was taken to avoid any participant consecutively participating in more than one progressive study or more than one resultative study.
vs. perfective), i.e., whether they were concurrent with or subsequent to the aspect activity. That is, the post-critical sentences did not indicate that the aspect activities were completed in the imperfective versions.

All the sentences used in the stories were simple sentences. No filler stories were included in this study to reduce participants’ fatigue, while filler probes were prepared to distract participants’ attention from the researcher’s intention. Table 6.1 illustrates the structure of a critical story.

Table 6.1. Structure of the critical stories

| Introduction | 3 sentences
| (that introduce the main characters and the context of the story) |
| Experimental | 1 critical sentence [followed by the near probe] |
| | 3 post-critical sentences [the far probe was presented immediately after one post-critical sentence] |
| Conclusion | 2 to 3 sentences (to conclude the story) |

Critical Sentences. The 24 test items were created by the researcher, using verbs that evoke progressive readings in the imperfective sentences for the Progressive study (Experiment 6), and using verbs that evoke resultative readings in the imperfective sentences for the Resultative study (Experiment 7). The critical sentences used a past progressive (V-tei-ta) or simple past tense (V-ta) for manipulation of the imperfective or perfective grammatical aspect. For the sake of more consistency across experiments, the verbs employed in Experiments 6 and 8 were from Experiments 3 and 4, respectively.\(^{35}\)

\(^{35}\) There was an exception of one item in each set of verbs that was changed between experiments: 聞く or 聴く ‘hear or listen to’ in progressive items, and 婚約する or 結婚する ‘get engaged or married’ in resultative items.
Experiment 6 (Progressive) used a proper name with a topic marker -wa for the subject noun to represent a single character. Experiment 7 (Resultative) used a proper name or a common noun for an inanimate object, with a topic marker -wa, as the subject noun in the resultative sentences is often a Patient/Theme that undergoes a change of state. The topic marker -wa was chosen over the nominative marker -ga for naturalness. Gender for the subject noun was counterbalanced such that female and male characters each appeared in half the stories.

In the stimuli that each participant saw, a quarter of the verb phrases had imperfective aspect followed by a near probe (17), a quarter had imperfective followed by a far probe (18), a quarter had perfective followed by a near probe (19), and a quarter had perfective followed by a far probe (20). The following example sentences illustrate the four conditions. Table 6.2 illustrates the positions where the probe appears. See Appendixes G and H for a complete list of critical sentences used in Experiments 6 and 7, respectively.

(17) Condition 1. Imperfective-Near: 沙織はココアを飲んでいた。
(18) Condition 2. Imperfective-Far: 沙織はココアを飲んでいた。
(19) Condition 3. Perfective-Near: 沙織はココアを飲んだ。
(20) Condition 4. Perfective-Far: 沙織はココアを飲んだ。

Gloss: Saori-Top cocoa-Acc {drink-tei-ta/ drink-ta}.
English Translation: ‘Saori was drinking/drank cocoa.’

Critical Probe: 飲む ‘drink’

Critical Probes. The 24 critical probes were created by using the untensed form (strictly speaking, the non-past form in Japanese) of the verbs used in the critical sentences. All 24 critical probes were intended to elicit a “yes” response, but the response latencies were
expected to be longer for weaker memory retrieval. No other elements (i.e., filler probes) besides the story sentences intervened between the critical sentence and the critical probe.

Table 6.2. Positions of critical probes for Experiments 6 and 7

<table>
<thead>
<tr>
<th>Critical Sentence</th>
<th>Post-Critical Sentence 1</th>
<th>Post-Critical Sentence 2</th>
<th>Post-Critical Sentence 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Saori was drinking/drank cocoa.</td>
<td>(She) made it thicker that day</td>
<td>(Her) body was getting warmed up from the core.</td>
<td>As (one) relaxes, (one) recovers her motivation.</td>
</tr>
</tbody>
</table>

**Positions for Critical Probes.** The critical probe was presented at either a near probe position or a far probe position during the presentation of each story. Following Magliano and Schleich’s (2000) procedure, the near probe was positioned immediately after the critical sentence in all four experiments (Experiments 6 to 9). In Experiments 6 and 7, the far probe was positioned immediately after the first sentence that followed the critical sentence. In Experiments 8 and 9, the far probe was positioned immediately after three sentences that followed the critical sentence.

Before the experiments were conducted, a norming study was conducted with a separate group of 11 participants to decide which probe position resulted in the largest effects. Of the three test locations for sentences (after the first, second, and third sentences following the critical sentence), the effects were observed to be largest at the far probe immediately after the first sentence that followed the critical sentence in stories with verbs that evoke resultative readings with the imperfective.
Magliano and Schleich (2000) placed the far probe three sentences after the critical sentence, since the effects were observed to be largest at this position for the far probe in the progressive use of the imperfective. This position was used for testing memory retrieval in the long span in Experiments 8 and 9.

*Filler Probes.* The filler probes were created and mixed into each story such that one to three filler probes in addition to the critical probe asked for recognition in each story. Filler probes refer to a sentence constituent at various sentence positions, and they were of various parts of speech such as adverbial, adjective, verb, and noun. The untensed, or non-past, form was used for verb probes. This secondary probe recognition task on filler probes served to distract participants’ attention from the researcher’s intention as well as to maintain and monitor participants’ attentiveness to the reading. Filler probes were intended to elicit either “yes” or “no” responses, which were counter-balanced so that the critical and filler probes together required approximately half “yes” and half “no” responses as the correct answer.

*Confirmation Questions.* A quarter of the 24 stories were followed by confirmation questions. These six questions were created separately for the Progressive study (Experiment 6) and the Resultative study (Experiment 7), by the researcher. They were simple yes/no questions asking about the content of the critical story. This secondary task served to maintain and monitor participants’ attentiveness to the reading. Half of all the question sentences required “yes” and the other half required “no” for the correct answer.
6.1.3 Procedure (Progressive, Resultative)

Participants read 24 short stories one sentence at a time, paced by their own key pressing. For one quarter of the 24 stories, they also answered a simple yes/no question that was presented immediately after a story, and that asked about the content of the story. The stories were presented in a pseudo-random order, such that participants encountered one story with a yes/no question per four stories. For each story, each sentence was presented by itself on the center of the computer screen. Participants were instructed to press the spacebar to read sentences, and to press the keys labeled はい‘yes’ or いいえ‘no’ to respond to the probes and the yes/no questions. Participants were instructed to read each sentence and answer each probe or question as quickly and accurately as possible, but to make sure to understand each sentence. In responding to a probe, participants were asked whether they thought they had read the word in preceding sentences in a given story. Participants were instructed to press the “yes” key if they thought they had, or the “no” key if they did not think they had. The “yes” key was positioned on the right and the “no” key on the left of the key board. Since the computer was measuring response times, participants were also asked to keep both hands on the keyboard during the test session to avoid any latency in key pressing unrelated to probe recognition. Each sentence was replaced by the next sentence, or a probe word, upon key pressing, and the last sentence of one-quarter of the stories was immediately followed by a question sentence upon the same key (i.e., the spacebar) being pressed. Question sentences were presented alone in the middle of the screen and in green, to signal a transition from the reading sentences presented in black. The question sentence stayed on the screen until the “yes” or “no” key was pressed by the participant, upon which a blank screen with a fixation cue appeared in
the middle of the screen, staying until participants pressed the spacebar to move on to the next story. A probe word was periodically presented on the center of the screen with letters in red. The probe word stayed on the screen until the participant pressed the “yes” or “no” key. Pressing either key brought the next sentence in the story or the blank screen with a fixation cue in the middle of the screen if the session was finished. A practice session preceded the main session with two practice stories. The instructions to participants were given visually and orally: Participants were instructed orally while seeing screens that presented written instructions. The presentation of stimuli in random order across items and participants, and the measurements were controlled and collected by E-prime software 2.0. The whole session took about 12 minutes.

6.1.4 Conditions (Progressive, Resultative)

Experiments 6 and 7 each tested four conditions across two variables: (i) grammatical aspect with two levels (Imperfective, Perfective), and (ii) probe position (Near, Far). Reaction times for accurate responses at the critical probes were the dependent measures in these experiments.

6.1.5 Predictions (Progressive, Resultative)

Experiments 6 and 7 tested whether memory retrieval for event representations constructed during on-line sentence processing may, or may not, be impacted by grammatical aspect alone, or by an interaction of lexical and grammatical aspect. They tested one particular hypothesis (i.e., Focusing Hypothesis) that such memory retrieval may, or may not, be impacted by changes in which part of an event is highlighted, using

166
Japanese imperfective sentences in which the interaction between lexical and grammatical aspect can change the highlighted event part. Experiments 6 and 7 particularly tested memory retrieval of event representation in a short span.

I predicted that the far probe would show longer reaction times than the near probe, assuming that memory decay is greater by the time the comprehender reaches the far probe position. Since the progressive use of the Japanese imperfective is equivalent to the English imperfective tested by Magliano and Schleich (2000), and since their study revealed that the imperfective helps memory retrieval of event representation more than the perfective, I predicted that the results from readers of the imperfective sentences in Experiment 6 (Progressive) would replicate Magliano and Schleich’s (2000) results, showing faster responses to the far probe in the imperfective than in the perfective. Since the role of the imperfective is to maintain information in working memory, the perfective and the imperfective conditions will have essentially the same level of activation at the near probe. The degree of effect in Experiment 6, which tests memory in a short span, may be smaller than the degree of effect in Experiment 8, which tests memory in a long span, since Magliano and Schleich’s results revealed an increased degree of effects on memory in a longer span than in a short span. No difference in results is predicted between Experiment 6 and Experiment 7 if the imperfective aspect uniformly plays a role in memory retrieval, since the expected effects originate in grammatical aspect alone, regardless of what verb type it accompanies and hence regardless of which part of the event it highlights (Uniform Hypothesis).

But if highlighting on event parts does relate to the roles for memory retrieval (Focusing Hypothesis), then I propose specific predictions as follows:
**Focusing Hypothesis.** Since the imperfective plays its role when it highlights the ongoing nature of an event versus its completion (as observed in English by Magliano and Schleich, 2000), readers of the imperfective sentences in Experiment 7 (Resultative) would respond no faster to the far probe in the imperfective than in the perfective, while in Experiment 6 (Progressive), responses to the far probe should be faster in the imperfective condition than in the perfective condition.

If there are any effects of the imperfective with end-state focus (Experiment 7, Resultative), they may be short-lived. To restate this, in this experiment, which tests memory retrieval in a short span, readers of the sentences in Experiment 7 (Resultative) might have faster responses to the far probe in the imperfective than in the perfective. Additionally, if memory retrieval is compared between a short span and a long span, any effects in memory retrieval in the short span (Experiment 7, Resultative) would be larger than those in the long span (Experiment 9, Resultative), assuming that the effects are short-lived.

6.1.6 Data analysis

6.1.6.1 Progressive (Experiment 6)

**Selected Participants, Items, and Response Times**

For Experiment 6, I report analyses of eight participants’ reaction times (RTs). The data cleaning was minimal, conducted at the time of the pilot analyses for deciding whether to
continue with the experiment.\textsuperscript{36} The average accuracy rate on the confirmation questions was 85.42%, and no participant was removed due to this criterion. The average accuracy rate on probe recognition on all critical and filler probes from all participants was 95%; no participant was removed due to probe recognition inaccuracy. As a result, data from eight people (i.e., two participants for each of four conditions) remained for subsequent analyses. In the next step, I replaced RTs as outliers if they were above a fixed cutoff point of 3000 milliseconds. The replacement rate of RTs by this outlier treatment in critical probes was 0.53%. Finally, the mean value and standard deviation of the probe response times were calculated for each participant from their correct responses, and a trial was replaced if its RT was below or above 3 standard deviations of the mean value. With this criterion, 2.82% of overall RTs on the entire set of trials was replaced.

\textbf{6.1.6.2 Resultative (Experiment 7)}

\textit{Selected Participants and Items}

First, the average accuracy rate on the confirmation questions from all participants was 99.6%; no participant was removed due to this criterion. Second, the participants’ accuracy rate on the probe recognitions was calculated based on the entire set of trials (including critical and filler probes), with the idea that participants should be excluded if their accuracy rate was below 80%. No participant was removed due to probe recognition.

\textsuperscript{36} Experiment 6 was not completed, as results of the initial analyses of the eight participants’ data were not promising. I considered it more reasonable to collect data in a longer-span experiment (i.e., my prospective Experiment 8, in which the far probe position will be further away than in Experiment 6): In Magliano and Schleich’s Experiment 1, they found the largest differences among their test locations 1 and 4 (with test location 1 being closest to the aspect sentence and location 4 the furthest); their imperfective grammatical aspect was in progressive readings (not resultative readings), as in my Experiment 6.
inaccuracy (the average accuracy rate from 40 people was 94%). Third, the item accuracy rate on the probe recognitions was calculated based on responses from all participants, with the idea that an item should be excluded if accuracy on the critical probe for the item was below 80%. No critical item was removed due to probe recognition inaccuracy (the average accuracy from the 24 critical probes was 95%). All remaining incorrect responses were filtered out for subsequent analyses. Fourth, the data were examined to identify outlying participants by calculating the mean value and standard deviation of response times of the entire set of trials (including critical and filler probes) for each participant, with the idea that participants should be excluded if their mean RT was below or above 3 standard deviations of the overall mean. No participant was removed due to the overall mean criteria (the average of overall mean RT from 40 people was 1044 milliseconds). Fifth, the data were examined to identify outlying items by calculating the mean value and standard deviation of response times to critical probes from all participants, with the idea that items should be excluded if their mean RT was below or above 3 standard deviations of the overall mean. No item was removed due to the overall mean criteria (the average of overall mean RT from the 24 critical probes was 1091 milliseconds). As a result, the data from all 40 people (i.e., 10 participants for each of four conditions) remained in the analysis. All incorrect responses were filtered out for subsequent analyses.

**Selected Response Times.** Outlying response times on the entire set of trials (including critical and filler probes) were trimmed by first replacing times that were over 5000 milliseconds with this cutoff value (10 trials, i.e., 0.32% of RTs were replaced due to this
threshold), with the idea that these times represented obvious outliers that would skew the subsequent mean value and standard deviation. Second, the new mean value and standard deviation of the probe response times were calculated for each participant. Trials were replaced with the cutoff value if their RT was below or above 3 standard deviations from that value. With this criterion, 3.40% of overall RTs on the entire set of trials was replaced.

6.1.7 Results

6.1.7.1 Progressive (Experiment 6)

Error rates and decision latencies were collected for the verb phrase recognition decisions. Error rates from the eight people on the critical probe were low, with a mean error rate of .059. The average accuracy rate on the confirmation questions from the eight participants was 85%, indicating that they were reading carefully. The results of Experiment 6, as a pilot study with a small number of participants, are summarized in Figure 6.1. First, a 2 (Aspect: Imperfective vs. Perfective) x 2 (Probe Position: Near vs. Far) two-factor analysis of variance (ANOVA) with aspect and probe position as repeated factors revealed a main effect of probe position only by subject (but it was robust: $F_1 (1, 7) = 12.8$, $p = 0.009$, $F_2 (1, 23) = 2.4988$, $p = 0.1276$), suggesting that the near probes were responded to faster (by participants, $M = 873.66$, $SE = 65.19$) than the far probes (by participants, $M = 1059.8$, $SE = 68.06$). No other significant differences were observed (all $Fs < 1.1$), implying that Japanese readers do not rely on grammatical aspect in
maintaining event representation in a situation model with the imperfective aspect
highlighting an ongoing part of an event.

Figure 6.1. Mean decision latencies in milliseconds at different probe positions from Experiment 6, by participants, with 95% confidence intervals for each condition group. Results are shown for two grammatical aspect conditions at two critical probe positions: Near probes (immediately after a critical sentence), and Far Probes (after one post-critical sentences).

<table>
<thead>
<tr>
<th>Probe Positions</th>
<th>Imperfective</th>
<th>Perfective</th>
</tr>
</thead>
<tbody>
<tr>
<td>Near</td>
<td>884</td>
<td>864</td>
</tr>
<tr>
<td>Far</td>
<td>1092</td>
<td>1028</td>
</tr>
</tbody>
</table>

6.1.7.2 Resultative (Experiment 7)

Error rates and decision latencies were collected for the verb phrase recognition decisions. The error rates from all 40 people on the critical probe were low, with a mean error rate of .049, indicating that they were attentive to the task. The average accuracy rate on the confirmation questions from the 40 participants was 99.6%, showing that they were reading carefully. Figure 6.2 summarizes the mean decision latencies. A 2 (Aspect: Imperfective vs. Perfective) x 2 (Probe Position: Near vs. Far) two-factor ANOVA with repeated measures on both factors was conducted on the recognition latencies at probe
recognition. It revealed a robust significant main effect of probe position both by participants ($F_1 (1, 39) = 62.1457, p < .0001$) and by items ($F_2 (1, 23) = 46.6253, p < .0001$), showing that the near probes were responded to faster (by participants, $M = 939.61, SE = 31.37$) than the far probes (by participants, $M = 1212.04, SE = 47.19$). However, there was no main effect of aspect nor was there an interaction between aspect and probe position either by participants or by items (all $F$s < 1, except for a main effect of aspect by a subject analysis, $F_1 (1, 39) = 1.5116, p = 0.226261$). Overall, the results indicate that Japanese readers do not rely on grammatical aspect in maintaining event representation in a situation model when either the imperfective or the perfective grammatical aspect highlights the end-state.

![Figure 6.2. Mean decision latencies in milliseconds at different probe positions from Experiment 7, by participants, with 95% confidence intervals for each condition group. Results are shown for two grammatical aspect conditions at two critical probe positions: Near probes (immediately after a critical sentence), and Far Probes (after one post-critical sentences).](image)
6.1.8 Discussion

The only significant effect on memory retrieval of event representation in short span memory was the main effect of probe position, as expected based on the claim made in Magliano and Schleich that memory decay is larger at a far probe position than a near probe position. The experiments’ results, however, counter the Focusing Hypothesis, because grammatical aspect had no impact on the accessibility of information in working memory during on-line Japanese sentence comprehension, and this was the case regardless of highlights on event parts (on-going part vs. end-state). No effects were observed even with the progressive reading of the imperfective aspect, which is equivalent to the condition for which Magliano and Schleich found effects of grammatical aspect. Therefore, the results of this test with Japanese seem inconsistent with the claim made by Magliano and Schleich that grammatical markers (e.g., grammatical aspect) provide processing instructions to the reader regarding the management of information in working memory (Carreiras et al., 1997; Garnham, Oakhill, Ehrlich, & Carreiras, 1995; Givón, 1992, 1995; Kintsch, 1992; Morrow, 1986; Sanford & Garrod, 1981). To summarize the findings, interaction of lexical and grammatical aspect was null, countering the Focusing Hypothesis. Crucially, it was strongly suggested that grammatical aspect itself does not have any impact on memory retrieval of event representation in Japanese sentence processing, although it does in English sentence processing (according to Magliano and Schleich).

It is unlikely that participant insensitivity to the factors being tested in these experiments could account for a lack of aspect effect. Good accuracy rates, both on probe recognition and responses to confirmation questions, indicate that the participants
attended to the task well. But the most important point is that the experiments found a main effect of probe position, telling us that the experiments did measure subtle activities in the readers’ sentence processing. Testing aspect effects in a shorter span than that used by Magliano and Schleich, the present experiments found that the accessibility of information in working memory is immediately affected by sentence position, even at a distance of only one sentence (one sentence being the difference between the Near Probe and the Far Probe).

The short distance, however, may explain the lack of effect of grammatical aspect. The items tested in these experiments might not have involved enough distance of the probes from the target aspect sentence for grammatical aspect to present its effect. The far probe of Experiments 6 and 7 was only one sentence apart from the target sentence, in contrast to the test in Magliano and Schleich, where the probe was three sentences after the target. In addition, the present Experiment 6 was not a complete study. There were different expectations for Experiments 6 and 7 (Progressive and Resultative, respectively), which both employed a far probe that was closer to the target sentence than in Magliano and Schleich’s study. According to Magliano and Schleich, who tested the progressive use of the imperfective aspect, we would be more likely to observe effects of a far probe at a more distant position. In contrast, Experiment 7 (Resultative) tested the imperfective with end-state focus, the effect of which may be short-lived; if so, any effect of aspect on the resultative reading should be stronger at a closer far probe. However, there was no effect. Experiment 6 (Progressive, short span) was not completed because the pilot analyses of eight participants’ data produced null results; I assume that, in this case, effects would be stronger at a further far probe position. I will be testing this possibility.
by conducting a long-span memory experiment (Experiment 8) on the aspect effect in Japanese sentences.

Crucially, I found null effects of grammatical aspect in the imperfective of resultative reading in this short span experiment (Experiment 7), counter to my expectation that any effects of the imperfective with end-state focus in this condition would be short-lived, which the pilot study preceding Experiment 7 indicated. If the degree of effects of the imperfective aspect in the resultative reading are found to be smaller in the prospective Experiment 9, which will test memory in a long span, than the effects in Experiment 7, which tested memory in a short span, these results will provide evidence of differences in the effect on reaction times of the closer versus the further positions.

It is worth noting that Magliano and Schleich suggested that the effect of grammatical aspect was not stable across their test groups. Magliano and Schleich conducted two similar probe recognition experiments (Experiments 3 and 4), in one of which they introduced a third factor, low versus high memory-span reader groups, using scores on a computerized working memory span test (from Kim, Millis, & Langston, 1999), along with the other two factors of grammatical aspect and probe position. In both experiments, the imperfective sentences induced faster responses than the perfective sentences at the far probe, but one subgroup of comprehenders who had relatively low-span memory showed null effects of grammatical aspect at both probe positions. The participants in the present experiments may be similar to the low-span memory participants in this sense. At this point, I save further discussion of this possibility until I have the results of Experiments 8 and 9.
6.2 Memory retrieval in long span

Experiments 8 and 9 were identical to Experiments 6 and 7, respectively, with the exception that they were focused on, and were conducted in, a long span condition for memory retrieval: The far probe in Experiments 8 and 9 was positioned immediately after three sentences after the critical sentence.

6.2.1 Participants (Progressive, Resultative)

Forty-one undergraduate students at Osaka University (in Toyonaka, Japan) participated in the Progressive study (Experiment 8), and a separate group of 33 undergraduate students at Osaka University participated in the Resultative study (Experiment 9), for compensation of $7 or $12. Participants were compensated $7 when they participated in this study alone, or $12 when they participated in either one of the present studies after they participated in another study (Experiments 3, 4, or 5) that took about five minutes. All were self-reported to be monolingual native speakers of Japanese.

6.2.2 Materials (Progressive, Resultative)

The materials that were used in Experiments 8 and 9 were identical to those that were used in Experiments 6 and 7, respectively, except that the far probe was positioned immediately after the three post-critical sentences in Experiments 8 and 9. Table 6.3 illustrates the positions where the probe appears. See Appendixes G and H for a complete list of critical sentences used in Experiments 8 and 9, respectively.

---

37Care was taken to avoid any participant consecutively participating in more than one progressive study or more than one resultative study.
Table 6.3. Positions of critical probes for Experiments 8 and 9

<table>
<thead>
<tr>
<th>Critical Sentence</th>
<th>Post-Critical Sentence 1</th>
<th>Post-Critical Sentence 2</th>
<th>Post-Critical Sentence 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Saori was drinking/drank cocoa.</td>
<td>(She) made it thicker that day</td>
<td>(Her) body was getting warmed up from the core.</td>
<td>As (one) relaxes, (one) recovers her motivation.</td>
</tr>
</tbody>
</table>

In addition, some filler probes that were previously positioned among the three post-critical sentences in Experiments 6 and 7 were given changed positions and referring words, so that no filler probes intervened between the critical sentence and the current far probe position.

6.2.3 Procedure (Progressive, Resultative)

The procedures used in Experiments 8 and 9 were identical to those that were used in Experiments 6 and 7.

6.2.4 Conditions (Progressive, Resultative)

The four conditions (2 grammatical aspects by 2 probe positions) tested in Experiments 8 and 9 were identical to those that were tested in Experiments 6 and 7. Dependent measures were response times for accurate responses to the critical probes.

6.2.5 Predictions (Progressive, Resultative)

Experiments 8 and 9 tested identical questions to those in Experiments 6 and 7, but in a long span for memory retrieval. The particular hypothesis that the memory retrieval may,
or may not, be impacted by changes in highlights in event parts has a separate prediction under *Focusing Hypothesis* in Experiments 8 and 9, from the predictions in Experiments 6 and 7.

The far probe will show longer reaction times, assuming that memory decay is greater at a far probe position than a near position. Since the long-span condition is identical to the condition that induced most of the effects of the imperfective in the study done by Magliano and Schleich (2000), and since the progressive use of the imperfective is equivalent to the English imperfective tested by Magliano and Schleich, I predicted that readers of the imperfective sentences in Experiment 8 (Progressive) would have faster responses to the far probe than they would with the perfective, since the imperfective activity potentially remains at a relatively high state of activation over subsequent context (Magliano & Schleich, 2000). The perfective and the imperfective activities will have essentially the same level of activation at the near probe. But the degree of effects at the far probe in the present experiment, which tests memory in a long span (Experiment 8, Progressive), may be larger than the degree of effects in Experiment 6 (Progressive), which tests memory in a short span, since Magliano and Schleich demonstrated an increased degree of effects in memory in a longer span than in a short span. If the imperfective aspect uniformly plays a role in memory retrieval, there should be no difference in results between Experiments 8 and 9, since the expected effects originate in grammatical aspect alone, regardless of what verb type it accompanies and hence regardless of which part of the event it highlights (*Uniform Hypothesis*).

But if highlighting on different parts of an event does relate to the roles for memory retrieval (*Focusing Hypothesis*), then I propose specific predictions as follows:
**Focusing Hypothesis**

Since the imperfective plays its role when it highlights the ongoing nature of an event versus its completion, as observed in English in Magliano and Schleich’s (2000) study, readers of the imperfective sentences in Experiment 9 (Resultative) should respond no faster to the far probe in the imperfective than in the perfective, while in Experiment 8 (Progressive) responses to the far probe should be faster in the imperfective condition than in the perfective condition.

Since memory decay has been shown to be faster with an end-state focus (perfective aspect) than with an ongoing state focus (imperfective) (Magliano & Schleich, 2000), any effects of the imperfective with end-state focus (Experiment 9, Resultative) should be short-lived. To restate this, in this experiment, which tests memory retrieval in a long span, I predict that readers of the imperfective sentences in Experiment 9 (Resultative) will have no faster response times to the far probe than they will with the perfective, since short-lived effects would fail to help memory retrieval in a long span. In other words, if compared in terms of a short or long span, reaction times to the far probe in a long span (Experiment 9, Resultative) would be slower than those in the short span (Experiment 7, Resultative), since the effects of the imperfective with the end-state focus are short-lived and thus would persevere better in a short span (Experiment 7, Resultative) than in a long span (Experiment 9, Resultative).

**6.2.6 Data analysis**
6.2.6.1 Progressive (Experiment 8)

*Selected Participants and Items.* First, the participants’ accuracy rate on probe recognition was calculated based on the entire set of trials (including critical and filler probes), with the idea that participants should be excluded if their accuracy rate was below 80%. No participant was removed due to probe recognition inaccuracy (the average accuracy rate from 41 people was 95%). Second, the items’ accuracy rate on probe recognition was calculated based on responses from all 41 participants, with the idea that items should be excluded if their accuracy on the critical probe was below 80% (recall that each of the 24 items contains one critical probe). Two critical items were removed due to low probe recognition accuracy (68% on item #12 歌う ‘sing’ and 73% on item #13 話す ‘talk’). The average accuracy for the remaining 22 critical items was 94%. All remaining incorrect responses were filtered out for subsequent analyses. Third, the data were examined for any outlying participants by calculating the mean value and standard deviation of response times (RTs) of the entire set of trials (including critical and filler probes) for each participant, with the idea that participants should be excluded if their mean RT was below or above 3 standard deviations of the overall mean. One participant was removed due to the overall mean criteria (the average of overall mean RTs from the remaining 40 people was 1131 milliseconds). Fourth, after the exclusion of the one participant, the data were examined for any outlying items by calculating the mean value and standard deviation of response times to critical probes from all 40 participants, with the idea that items should be excluded if their mean RT was below or above 3 standard deviations of the overall mean.

38 Because item #13 in the critical sentence contained 鼻歌を歌う ‘hum’, participants might have been misled about whether prior sentences contained 歌う ‘sing’, although the lexical item (歌う) was identical in Japanese.

39 The results were not changed by removing these two items for the analyses.
above 3 standard deviations of the overall mean. No item was removed due to the overall mean criteria (the average of overall mean RT from 22 critical probes was 1159 milliseconds). Fifth, in order to have an equal number of participants for each condition, four more participants (those with the lowest accuracy rates on the probe recognitions) were removed. As a result, 36 people’s data (i.e., nine participants for each of four conditions) remained for subsequent analyses. All incorrect responses were filtered out for subsequent analyses.

**Selected Response Times.** Outlying response times were trimmed by first replacing times that were over 5000 milliseconds with the cutoff value (14 trials, or 0.51% of RTs, were replaced due to this threshold) on the entire set of trials (including critical and filler probes), with the idea that these times represented obvious outliers that would skew the subsequent mean value and standard deviation. Second, the new mean value and standard deviation of the probe response times were calculated for each participant. Trials were replaced if their RT was below or above 3 standard deviations from the mean value. With this criterion, 0.94% of overall RTs on the entire set of trials were replaced.

**6.2.6.2 Resultative (Experiment 9)**

**Selected Participants and Items.** First, the participants’ accuracy rates on the probe recognition were calculated based on the entire set of trials (including critical and filler probes), with the idea that participants should be excluded if their accuracy rate was below 80%. No participant was removed due to probe recognition inaccuracy (the average accuracy rate from 33 people was 95%). Second, the items’ accuracy rate on the
probe recognition was calculated based on responses from all 33 participants, with the idea that items should be excluded if their accuracy on the critical probe was below 79% (recall that each of the 24 items contains one critical probe). Two critical items were removed due to low probe recognition accuracy (67% for item #7 つける ‘wear’ and 76% for item #12 止まる ‘stop’). The average accuracy from the remaining 22 critical items was 90%. All remaining incorrect responses were filtered out for subsequent analyses. Third, the data were examined for any outlying participants by calculating the mean value and standard deviation of response times for the entire set of trials (including critical and filler probes) for each participant, with the idea that participants should be excluded if their mean RT was below or above 3 standard deviations of the overall mean. One participant was removed due to the overall mean criterion (the average of overall mean RTs from the remaining 32 people was 1078 milliseconds). Fourth, after the exclusion of the one participant, the data were examined for any outlying items by calculating the mean value and standard deviation of response times to critical probes from all 32 participants, with the idea that items should be excluded if their mean RT was below or above 3 standard deviations of the overall mean. No item was removed due to the overall mean criterion (the average overall mean RT from 22 critical probes was 1153 milliseconds). Fifth, in order to have an equal number of participants for each condition, four more participants (those with the lowest accuracy rates on probe recognition) were removed. As a result, 28 people’s data (i.e., seven participants for each of four conditions) remained for subsequent analyses. All incorrect responses were filtered out for subsequent analyses.

40 The results were not changed by removing these two items in the analyses.
Selected Response Times. Outlying response times were trimmed by first replacing times that were over 5000 milliseconds with the cutoff value on the entire set of trials (including critical and filler probes), with the idea that these times represented obvious outliers that would skew the subsequent mean value and standard deviation (7 trials, or 0.33% of RTs, were replaced due to this threshold). Second, the new mean value and standard deviation of the probe response times were calculated for each participant. Trials were replaced if their RT was below or above the 3 standard deviations of the value. With this criterion, 3.87% of overall RTs on the entire set of trials were replaced.

6.2.7 Results

6.2.7.1 Progressive (Experiment 8)
Error rates and decision latencies were collected for the verb phrase recognition decisions. The error rates from all 36 people on the critical probe were low, with a mean error rate of .058, indicating that they were attentive to the task. The average accuracy rate on the confirmation questions from the 36 participants was 85%, showing that they were reading carefully. Figure 6.3 summarizes the mean decision latencies. A 2 (Aspect: Imperfective vs. Perfective) x 2 (Probe Position: Near vs. Far) within subjects repeated ANOVA (analysis of variance) was conducted on the decision latencies. It revealed a significant main effect of probe position both by participants ($F_1(1, 35) = 56.1924, p < .0001$) and by items ($F_2(1, 21) = 90.5652, p < .0001$), showing that the near probes were responded to faster (by participants, $M = 911.21, SE = 36.61$) than the far probes (by participants, $M =$
However, there was no significant main effect of aspect, nor a significant interaction of aspect and probe position by either participants or items (all $F$s < 1.2). The null aspect results from this experiment indicate that Japanese readers do not rely on grammatical aspect in maintaining an event representation in long-span memory for a situation model with an ongoing part of the event highlighted by the imperfective.

Figure 6.3. Mean decision latencies in milliseconds at different probe positions from Experiment 8, by participants, with 95% confidence intervals for each condition group. Results are shown for two grammatical aspect conditions at two critical probe positions: Near probes (immediately after a critical sentence), and Far Probes (after three post-critical sentences).

<table>
<thead>
<tr>
<th>Probe Positions</th>
<th>Near</th>
<th>Far</th>
</tr>
</thead>
<tbody>
<tr>
<td>Imperfective</td>
<td>913</td>
<td>1325</td>
</tr>
<tr>
<td>Perfective</td>
<td>910</td>
<td>1266</td>
</tr>
</tbody>
</table>

6.2.7.2 Resultative (Experiment 9)

Error rates and decision latencies were collected for the verb phrase recognition decisions. Error rates from the 28 people on the critical probe were low, with a mean error rate of .078. The average accuracy rate on the confirmation questions from all participants was 99%, indicating that they were reading carefully. Figure 6.4 summarizes the mean
decision latencies. A 2 (Aspect: Imperfective vs. Perfective) x 2 (Probe Position: Near vs. Far) within subjects repeated measures ANOVA was conducted on the recognition latencies. It revealed a significant main effect of probe position both by participants ($F_1(1, 27) = 36.4278, p < .0001$) and by items ($F_2(1, 21) = 108.6322, p < .0001$), showing that the near probes were responded to faster (by participants $M = 908.01, SE = 35.85$) than the far probes (by participants $M = 1351.35, SE = 87.42$). However, there was no significant main effect of aspect or significant interaction of aspect and probe position by either participants or items (all $F$s < 1), indicating that Japanese readers do not rely on grammatical aspect in maintaining event representation in long span memory for a situation model with the imperfective focusing on the end-state of the event.

Figure 6.4. Mean decision latencies in milliseconds at different probe positions from Experiment 9, by participants, with 95% confidence intervals for each condition group. Results are shown for two grammatical aspect conditions at two critical probe positions: Near probes (immediately after a critical sentence), and Far Probes (after three post-critical sentences).
6.2.8 Discussion

Testing the effect of aspect in long-span memory retrieval, Experiments 8 and 9 had different predictions than Experiments 6 and 7, which tested effects in short-span memory. Experiments 6 and 7 compared response latencies at two probe positions: at a position immediately after the aspect sentence (the near probe), and at a position that followed a post-aspect sentence (the far probe). In contrast, in Experiments 8 and 9 the far probe was positioned two sentences further away. This configuration demands more memory resources. First, Experiment 8 tested the effect of the progressive aspect with test materials in which the critical sentence for both near and far probes was at the same distance as the critical sentences used by Magliano and Schleich (2000) to test the progressive use. Therefore, I predicted that, as Magliano and Schleich observed for English, the responses to the far probe in Experiment 8 would be faster in the imperfective condition than in the perfective condition. However, this prediction was not borne out: Readers of the perfective sentences responded no more slowly to either probe than readers of the imperfective sentences. Second, the predictions for Experiment 9 differed from the predictions for Experiment 7, especially in regard to the Focusing Hypothesis, because of the greater probe distance in the current study. Because short-lived effects would fail to help memory retrieval in a long span, the prediction for Experiment 9 was that readers would not have faster response times to the far probe in imperfective sentences than in perfective sentences. The results confirmed this null effect of grammatical aspect, apparently conforming to the Focusing Hypothesis. However, this hypothesis would also predict positive effects of aspect with the progressive imperfective
(Experiment 8) and with the resultative imperfective in a short span condition (Experiment 7), neither of which were observed. Therefore, the complete set of experiments does not strongly support the Focusing Hypothesis.

The Uniform Hypothesis was not supported, either. As this hypothesis would predict, Experiments 8 and 9 showed no differences in any effect of grammatical aspect. The null results seem to indicate that the role of the imperfective aspect in memory retrieval does not vary depending on verb type. However, the imperfective aspect did not play a role in memory retrieval in the first place. Hence, the experiments do not support the Uniform Hypothesis.

The only significant effect that was observed was the robust effect of probe position in Experiments 8 and 9, which positioned the far probe in the longer span. The far probe induced slower responses than the near probe, indicating that memory decay is greater at the far probe position than the near position, as also observed in Magliano and Schleich’s experiments. The high accuracy rates on the probe recognition and confirmation questions suggest that it is very unlikely that the null effects of grammatical aspect reflected readers’ inattentiveness to the experiment.

To summarize the experiments testing aspect effects in long span memory, the Japanese readers did not show any evidence of using the information from grammatical aspect to maintain an event representation in their memory. The next section discusses the reasons for the contrast between the findings of this study on Japanese and the findings of the study on English conducted by Magliano and Schleich.
6.3 General discussion

The set of four Japanese experiments on which this chapter reports demonstrated that grammatical aspect had no perceptible impact on maintaining event representations in working memory in Japanese. This null effect of aspect is seemingly inconsistent with the claim that grammatical markers, such as aspect, provide processing instructions to the reader regarding the management of information in working memory (Carreiras et al., 1997; Garnham et al., 1995; Givón, 1992, 1995; Kintsch, 1992; Morrow, 1986; Sanford & Garrod, 1981). The results suggest that the memory representation of events constructed during on-line sentence processing was not impacted by grammatical aspect alone, nor by interaction of aspect and other factors (such as probe position, and/or progressive vs. resultative readings of the imperfective). For instance, the experiments introduced two different far probe positions to the manipulation of probe positions (the near vs. the far probe) employed by Magliano and Schleich (2000). This approach was taken because I was interested in the different predictions for this factor that relate to the different readings (progressive vs. resultative) of the Japanese imperfective, which depend on verb type: the imperfective aspect highlights the on-going parts of the event in Experiments 6 and 8 (progressive) but highlights the end-state in Experiments 7 and 9 (resultative). Because the effect of the imperfective is the maintenance of information in working memory, the perfective and the imperfective conditions will have essentially the same level of activation at the near probe. But at the far probe, grammatical aspect would be expected to manifest its impact in different ways for progressive and resultative imperfectives. That is, we would expect to see stronger effects at the three-sentence-distant far probe position in the progressive reading, but at the one-sentence-distant far
probe position in the resultative reading, because the effects, if any, should be short-lived with the end-state focus. However, the results indicated that grammatical aspect had no interaction with probe position, or progressive vs. resultative readings of the imperfective, casting doubt on the Focusing Hypothesis. The uniformly null effects seem to conform to the Uniform Hypothesis, but this hypothesis gained no support because the effects were null in the first place. The only effect, which was consistent across all the experiments, was a main effect of probe position, as in Magliano and Schleich’s study, confirming that early probes were responded to faster than far probes. Overall, this study did not replicate Magliano and Schleich’s findings that the activation level for an imperfective activity decayed at a slower rate than that for a perfective activity. Furthermore, their claim that readers maintain imperfective activities and events as ongoing in the situation models in subsequent contexts was apparently not supported in Japanese. These null results even appeared in Experiment 7, which was equivalent to Magliano and Schleich’s experiment as it used progressive imperfective at the same probe distance.

These four Japanese experiments, with their unexpected null effects of grammatical aspect, failed to support the Focusing Hypothesis. However, the predictions based on this hypothesis might be tested by directly comparing the four experiments to examine whether the expected effects appear. For instance, the hypothesis suggests the following predictions: (A) In the progressive use (Experiments 6 & 8), the effect of grammatical aspect will be larger at the far probe in a long span (Experiment 8) than in a short span (Experiment 6). On the other hand, (B) in the resultative use (Experiments 7 & 9), the effect of grammatical aspect will be smaller at the far probe in a long span (Experiment 9) than in a short span (Experiment 7). In addition, (C) within a short span
(Experiments 6 & 7), the effects of grammatical aspect at the far probe might be no
different between the progressive (Experiment 6) and the resultative (Experiment 7). On
the other hand, (D) within a long span (Experiments 8 & 9), the effects of grammatical
aspect at the far probe might be greater in the progressive (Experiment 8) than in the
resultative (Experiment 9). Although the results of such direct comparisons may not be
entirely straightforward, considering that the lexical items and subsequent sentences are
different between the progressive and resultative experiments, statistical analyses were
conducted to assess the hypothesis. One motivation for conducting these analyses was
that, numerically, the reaction times to the far probe position were greater in the two
resultative studies (Experiments 7 and 9) than in their progressive counterparts. No
statistical testing for (A) and (C) was conducted, because Experiment 6 (Progressive,
short span) lacked a large enough sample size (i.e., only eight participants).

(B): As has been discussed throughout this chapter, memory decay has been
reported to be faster with the end-state focus than the ongoing-state focus, and thus any
effects of the resultative imperfective on maintaining an event representation in memory
should be short-lived. The response latencies in Experiments 7 (short-span; far probe
after one sentence) and 9 (long span; far probe after three sentences) were compared to
test for the expected effects of the different far probe positions on reaction times. A two-
factor mixed-measures ANOVA with one factor (Aspect: Imperfective vs. Perfective) as
the repeated measures (i.e., within-subjects) factor and with another factor (Far probe
position: short-span vs. long-span) as the between-subjects factor was conducted on the
recognition latencies at the far probe.\textsuperscript{41} The analysis found an approaching main effect of Far probe position by items\textsuperscript{42} ($F_1(1, 45) = 3.68, p = 0.061572$) suggesting that the far probes were responded to faster in a short span ($M= 1215.25, SE= 45.40$) than in a long span ($M = 1340.42, SE = 46.87$) but not by subjects ($F_2(1, 67) = 2.28, p = 0.135825$), indicating no support for the idea that event information is more accessible in memory at a closer far probe position than a further probe position. There was no main effect of aspect, and no interaction of aspect and Far probe position either by participants or by items (all $F$s $< 1$, except for a main effect of aspect by a subject analysis, $F_1(1, 68) = 1.36, p = 0.247734$).

(D): Experiments 8 and 9 focused on manipulating the part of an event that was highlighted in the imperfective sentences to test whether a focus on the ongoing part versus the endpoint would affect the rate of decay of activation of an event representation over subsequent sentences. Experiment 8 used verbs that evoke a progressive reading in the imperfective, while Experiment 9 used verbs that evoke a resultative reading in the imperfective. In these experiments, the highlighted event parts were compatible with either the ongoing nature of the event or with the end-state of the event. Because memory decay has been shown to be faster with the end-state focus (in the perfective aspect) than with the ongoing-state focus (in the imperfective aspect) as discussed by Magliano and Schleich (2000), response latencies in the resultative-reading imperfective sentences (Experiment 9) were expected to be longer than those in the progressive-reading

\textsuperscript{41}The equivalent ANOVA on response latencies at the near probe showed no effect of aspect, probe position, or interaction by subjects (All $F$s $< 1$) or by items (all $F$s $< 1$ except for probe position, $F_1(1, 45) = 1.24, p = 0.271517$).

\textsuperscript{42}Due to the removal of two items from Experiment 9, the number of items in Experiments 7 and 9 were not equal (24 and 22, respectively).
imperfective sentences (i.e., Experiment 8) at the far probe. This possibility was assessed by comparing the response latencies at the far probe in Experiments 8 and 9. The analysis found no statistical differences between the two experiments’ response latencies at the far probe. A two-factor mixed-measures ANOVA with one factor (Aspect: Imperfective vs. Perfective) as the repeated measures (i.e., within-subjects) factor and with another factor (Reading: Progressive vs. Resultative) as the between-subjects factor did not show any effect of aspect or reading, or any interaction between aspect and reading by either participants or items (all $F$s < 1, except for aspect by subject analysis, $F_{1}(1, 64) = 1.17, p = 0.28$), indicating that the highlighting of event parts did not affect the role of grammatical aspect in maintaining representation in memory at a further position in subsequent sentences.

Against the Focusing Hypothesis, then, I conclude from this additional statistical analysis that the experiments provide no evidence that grammatical aspect plays a role in memory retrieval of event representation at any position in subsequent sentences, regardless of what part of a given event is highlighted.

We now turn to the question of what processing mechanism may account for why we see no effect of grammatical aspect in Japanese sentence processing. There are at least three possibilities to explain the present null effects of grammatical aspect: (i) the participants were mostly low working-memory span readers; (ii) Japanese readers are (apparently) all low working-memory span readers; and/or (iii) grammatical aspect has no impact on memory retrieval in Japanese sentence comprehension.

---

43 The equivalent ANOVA on response latencies at the near probe showed no effect of aspect or reading, nor any interaction by subject or by items (all $F$s < 1).
The first possibility is that the participants were all low working-memory span readers. As stated in Section 6.1.8, it is worth noting that the null aspect effects observed in the current study are similar to those observed in Magliano and Schleich’s study only for their participants with low working-memory spans. Because maintaining representations in memory is memory-consuming, Magliano and Schleich argued that the low working-memory span readers might lack working-memory resources to take advantage of aspect in order to maintain imperfective activities in memory. Considering that the current study did not measure its participants’ working-memory span, as Magliano and Schleich did in their Experiment 4, a future experiment should confirm the present results with a group of high working-memory span readers of Japanese, if any, who would have more resources for situation model construction than low working-memory span readers (Just & Carpenter, 1992). Alternatively, Japanese sentences, for some reason, may by default create too great a working memory load for readers to take advantage of aspect, leading to apparent null effects. This idea would not, however, account for the lack of difference between imperfective and perfective sentences, if processing grammatical aspect itself demands too much memory.

As discussed in Section 5.4, our account must therefore consider how the role of grammatical aspect might be better discussed by considering its availability (i.e., timing) in sentence processing. Specifically, it is a matter of how and when grammatical aspect plays a role in the construction of situation models. I interpret the present results to imply that the extent to which an event representation stays resilient is grounded in how early in the process of constructing situation models the information from grammatical aspect is

---

44 In their Experiment 3, Magliano and Schleich did not employ this factor and they did observe an aspect effect.
available to readers. This idea of resiliency of event representation has not clearly been
discussed from this perspective in previous studies (e.g., Magliano & Schleich). The
claim made in the literature has been that grammatical markers provide important
processing instructions regarding how to construct a situation (Carreiras et al., 1997;
Garnham et al., 1995; Gernsbacher, 1990; Gernsbacher & Jescheniak, 1995; Gernsbacher
Sanford & Garrod, 1981; Sanford et al., 1988; Zwaan, 1996). But assuming that sentence
processing is incremental (e.g., Altmann & Steedman, 1988; MacDonald, Pearlmutter, &
Seidenberg, 1994), situation models would be constructed part by part, and much of a
situation model would already have been constructed before the sentence-final position.
For Japanese readers, this is the position where they would first receive any instructions
from grammatical aspect. Therefore, the impact of aspect on resiliency in memory could
be non-existent with Japanese sentences, where imperfective events that are supposed to
be constructed on an ongoing basis in situation models may be no more measurably
resilient than perfective events.

According to Magliano and Schleich (2000), grammatical aspect provides cues
for what events will continue in the subsequent sentences as in cataphoric devices
(Gernsbacher & Jescheniak, 1995) rather than directs readers backwards for coreference
as in anaphoric devices (Garnham, Oakhill, Ehrlich, & Carreiras, 1995). While this
statement may be true for a language such as English that has grammatical aspect
available relatively early in a sentence, it might not apply to Japanese, because Japanese
is a verb-final (and aspect-final) language. Cataphoric devices tag the incoming
information so that the information is available in the subsequent context. Gernsbacher
and Jescheniak (1995) showed that the impact on memory representation of grammar-based cataphoric devices, such as an indefinite *this*, is more resilient in the subsequent context than the impact of anaphoric devices. In Japanese sentences, there comes no such information even if grammatical aspect would tag the “incoming” information. As discussed in Section 5.4, mental models are constructed on a clause basis, and models are monitored for any updates to be made on them as the reader moves onto the next clause (for discussion of mental models and time shift between clauses, see Zwaan, Langston, and Graesser, 1995). Japanese may well direct readers backwards for the construction of situation models, like anaphoric devices. Grammar-based markers of anaphor are reported to be short-lived and last about as long as the surface form of a text is maintained in working memory (Garnham et al., 1995). As Magliano and Schleich (2000, p.109) state, “To take advantage of grammar-based cues of anaphor, readers must consult their memory representation for surface level information.”

This study’s null aspect effect also implies that the role of grammatical aspect in relation to the iconicity assumption might be better discussed separately from the question of event representation resiliency in memory. The iconicity assumption represents the idea that readers assume that events and activities described in language occur in the order of description, successively on a timeline (Dowty, 1986; Fleischman, 1990), and grammatical aspect guides readers as to whether this iconicity assumption should be followed or not. The imperfective aspect cues a reader that the activity is not completed, and that the following activity should concur with the imperfective activity. Magliano and Schleich (p. 94) claim that “[the following activity] should therefore be temporally attached to [the imperfective activity] in a situation model
However, the present study’s results indicate no more impact of imperfective aspect on a situation model than of perfective aspect. Future experiments should thus explicitly assess Japanese readers’ perception of the duration of narrative activities on a timeline. Such an assessment could be carried out by having yes-no questions at near and far probe positions, as in Magliano and Schleich’s Experiments 1 and 2, in which they asked their participants whether they considered the activity to be still ongoing on the narrative timeline or not. The dependent measure would be the probability that the aspect activity was still considered ongoing. A finding that the imperfective activities were more likely to be considered ongoing than the perfective activities would indicate that Japanese readers use grammatical aspect for a cue on whether to follow the iconicity assumption in a way that is not reflected in activation levels of event representations.

Additional further experiments should re-confirm the impact of grammatical aspect in relation to the timing of its availability in sentence processing. I propose to do so by employing English sentences that have grammatical aspect sentence-finally, such as *John {walked/was walking}*. Because such English sentences have only the subject and the verb/aspect, I suggest adding a modifier to the subject noun, as in *The cat that John had been keeping for a while {walked/was walking}* in order to make grammatical aspect available relatively further down in a sentence. As grammatical aspect is expected to have little impact on the construction of situation models when it appears sentence-finally, one prediction is that the advantage of imperfective sentences for maintaining representations in memory would be reduced even in English sentence processing, conforming to the results from this study with Japanese sentences.
Finally, the current study’s null effects of grammatical aspect might be taken to suggest additional support for a claim made by Magliano and Schleich, which is that grammatical aspect plays a role in *maintaining* the activation level of information rather than *enhancing* it. If the imperfective, by marking information, increases the activation level more than the perfective (e.g., by tagging the information to make it more accessible to working memory), such an effect may be no different where grammatical aspect appears earlier in a sentence and where it appears later in a sentence. On the other hand, the timing of the availability of grammatical aspect in a sentence would have an impact on its effect if the activation level of event representation remains high as the result of grammatical aspect having been much used by the comprehenders for their enriched construction of situation models.

The present study, unlike any previous study, has assessed the extent to which grammatical aspect has an impact on the activation level of information and its accessibility to working memory and shown that the behavior of readers processing Japanese sentences is consistently different in this regard from the behavior of readers processing English sentences, suggesting the importance of taking the timing of grammatical aspect’s availability in processing into account when the role of grammatical aspect is discussed.
CHAPTER 7

GENERAL DISCUSSION

7.1 Original claims and new claim

In its exploration of human sentence comprehension, the present study has focused on how we comprehend time in events. It has extended the research on grammatical aspect with specific attention to the sources of the temporal properties that appear in a sentence besides grammatical aspect. The study began with the claim that grammatical aspect manifests its effect in interaction with other cues of temporal properties, particularly lexical aspect, which gives core information of a given event. This assumption of an interaction effect led to the predictions that (i) mismatching combinations of lexical and grammatical aspect would interfere with grammatical aspect’s effect, (ii) processing costs would reflect world knowledge of an event—specifically, its duration—in interaction with the effect of grammatical aspect, and (iii) the prominence of an event part highlighted by grammatical aspect (i.e., on ongoing part or end-state) would interact with the effect of grammatical aspect in memory retrieval. The study utilized Japanese as a test language because its imperfective grammatical aspect in combination with the lexical aspect of certain verbs (achievement verbs) creates a unique situation in which prominence on the end-state of an event co-occurs with a resultative reading, rather than the imperfective highlighting an ongoing, uncompleted part of an event with a progressive reading, which is the more common situation that has been widely explored in many previous studies on grammatical aspect.
However, the findings from a series of experiments with English and/or Japanese provided no support for these original claims. The results of the experiments testing “mismatching” combinations of lexical and grammatical aspect conformed to the aspectual coercion explanation. Furthermore, the experiments found no processing correlates of interaction between grammatical aspect and event duration or between grammatical aspect and verb types. In more specific terms, (i) the participants interpreted a described situation as completed when the perfective aspect was used even with an event without an end point (e.g., John kicked balls to Mary), suggesting that they made the coerced interpretation where an event was completed; (ii) the experiments controlling event duration using numerals (e.g., for two hours, golf three holes, etc.) found no interaction with grammatical aspect in reading times (in both English and Japanese); and (iii) the Japanese experiments provided no opportunity to obtain evidence for an interaction that could induce a change of highlighting on event parts, because the experiments quite unexpectedly found null effects; that is, they found no evidence of processing correlates of grammatical aspect itself at all in Japanese sentence processing.

The present findings, therefore, undermine the original claims but suggest a new claim: The manifestation of grammatical aspect’s role and effects in situation model construction during sentence processing is not interfered with by interaction with lexical aspect; but at the same time, its psychological reality in human sentence processing is not as universally stable as assumed in previous studies. In this regard, and given that processing studies that focus on the effects of grammatical aspect are still quite limited in number, more research is needed before we can reach any final conclusions.
7.2 Implications and ramifications

This study’s results conform to the idea that grammatical aspect functions as the “viewpoint” aspect even when there is no compatible event part for grammatical aspect in a given event structure. Specifically, the study observed that the perfective aspect places an outside viewpoint on an event that has no end-point (such as kick balls). Traditionally, aspectual coercion has been discussed for cases such as The light flashed until dawn, which shows a coerced interpretation due to the inconsistency of temporal properties between the verb phrase (e.g., the light flashed) and the modifying phrase (until dawn).

The present study provides evidence that such aspectual coercion extends to inconsistency between lexical and grammatical aspect. To my knowledge, this is the first study to suggest that aspectual coercion may occur at this level of inconsistency. Although prior studies have indicated that such inconsistency does affect sentence processing speed (Yap et al., 2006, 2009), the present study’s results suggest that a mismatch with lexical aspect does not interfere with grammatical aspect’s effect.

On the other hand, this study also showed that this function of grammatical aspect is not necessarily reflected in processing correlates, such as reading times or reaction times in a task like word recognition. In prior studies such as those by Anderson, Matlock, and Spivey (2010) and Magliano and Schleih (2000), which showed such effects of grammatical aspect, the experimental tasks were implemented in such a way that the contextual information was fairly available for a given event. For example, Anderson et al. provided a context of hard versus easy terrain before their participants heard their critical aspect sentence that described a motion event on the terrain. Magliano and Schleich (2000) used different verbs in comparing event durations, using the nature of the
events themselves (e.g., watching a movie for a long-durational event and scratching your nose for a short-durational event), in contrast to the present study, which used numerals for the contrast. In addition, Becker, Ferretti, and Madden-Lombardi (2013) revealed an effect of time shift between events by manipulating the duration of an intervening event, and they chose to use “events that inherently vary in duration rather than using the same events and modifying their duration with temporal adverbs” (p. 219). As Becker et al. reflected, using different events in the two conditions may have contributed to the observed differences, instead of the differences being based on pure event duration. In light of my argument, using different events may have contributed to more concrete construction of situation models that may have internalized event duration more strongly in mental representations. Some researchers claim that language comprehension is partial: It may not be complete, detailed, or accurate, as traditional theories assumed (see Ferreira, Bailey, & Ferraro, 2002). For instance, people can easily be misled by garden-path sentences (e.g., While John hunted the frightened deer escaped; Frazier, Carminati, Cook, Majewski, & Rayner, 2006) or reduced relative clauses (e.g., The performer sent the flowers was pleased; Rayner, Carlson, & Frazier, 1983) and not necessarily obtain a perfect interpretation from the beginning. As discussed by Ferreira et al. (2002), comprehenders may have merely “good enough” representations for sentence comprehension. Pickering, McElree, Frisson, Chen, and Traxler (2006) suggested that comprehenders do not immediately need to commit fully during normal reading; full commitment may occur only when processing demands induce immediate decisions. In their experiments, the aspectual coercion required in sentences such as The insect hopped effortlessly until it reached the garden was proved not costly in a self-paced reading task
and eye-tracking experiment in normal reading, although prior studies (Piñango, Winnick, Ullah, & Zurif, 2006; Piñango, Zurif, & Jackendoff, 1999; Todorova, Straub, Badecker, & Frank, 2000) claimed that such sentences should be costly because of the additional semantic computation of coercion. Pickering et al.’s claim was that the previous findings were the product of sentence comprehension where comprehenders were made attentive enough to the sentence by task demands. The present study predicted that long-duration events in imperfective sentences would have more processing cost than short-duration events, but observed no such interaction. My interpretation of the present results is that psychological activities depend on how deeply grammatical aspect participates in the construction of situation models. Information of events from different lexical items adds more concreteness to described situations in sentence comprehension than the numerically controlled difference of event duration in the present study, which may not have contributed much to the construction of situation models.

This study’s claim was that greater involvement of grammatical aspect in the construction of situation models would be reflected in more processing activity, but this was not supported by the (null) results from the Japanese experiments. I examined processing costs imposed by grammatical aspect (in an attempt to replicate Coll-Florit & Gennari, 2011) and by interaction between grammatical aspect and event duration in a phrase-by-phrase self-paced reading task. In addition, the role of grammatical aspect in making the event representation accessible longer in memory, either alone or in interaction with lexical aspect, was tested in Japanese (in an attempt to replicate Magliano & Schleich, 2000). Counter to the processing behaviors attested in the studies on English and/or Spanish, where they were explained as evidence of readers’
construction of durative situations (Coll-Florit & Gennari 2011), or of readers having the imperfective activities ongoing in a situation model (Magliano & Schleich, 2000: “imperfective activity potentially remains at a relatively high state of activation over subsequent context because it is ongoing in a situation model”), this study’s Japanese experiments found null processing correlates.

Null processing correlates that reflect real-time processing activities may not necessarily indicate that there is no construction of situation models. The fact that no processing correlates of grammatical aspect were detected in Japanese in processing times or in memory resiliency does not necessarily argue against the role of grammatical aspect. Because people interpret sentences incrementally (e.g., Marslen-Wilson, 1973), they do not wait until the end of a clause or sentence to start comprehending a given sentence (e.g., Fodor, Bever, & Garrett, 1974). Readers of Japanese imperfective or perfective sentences start their process of sentence comprehension before they read the grammatical aspect at the end of the clause or sentence, and importantly, those unfolding contents are identical until grammatical aspect finally appears. It is hence natural that grammatical aspect has minimal (or even null) effects in sentence processing in Japanese sentences, because the grammatical aspect only becomes available at the end of a sentence, rather than earlier in a sentence as in English or Spanish. In a study on Japanese, Ueno and Kehler (2010) replicated studies on English by Rohde, Kehler, and Elman (2006), Rohde and Kehler (2008), and Ferretti et al. (2009). The Japanese speakers showed less goal bias after reading imperfective sentences than after reading perfective sentences (e.g., Taro handed/was handing a book to Jiro. He___), indicating that the constraint of the perfective grammatical aspect to refer to the end-state of an event was
loosened in the imperfective sentences in Japanese, as observed in the English studies. In their study, the dependent measure was the likelihood of choosing the goal versus the source entity after reading the critical aspect sentence, rather than processing activity correlates such as reading times. If the effect of grammatical aspect on processing activities seems weak in Japanese, it may be because the information of grammatical aspect is available only sentence-finally and it may have less impact on the entire process of the situation model construction. However, the limitations of self-paced reading tasks may also be relevant, especially when there is no spill-over region, as in the case of the present study. Because the critical region of grammatical aspect occurs sentence-finally according to the word order in Japanese, we miss the chance to observe any spillover effects from the previous critical region at the successive regions.

Some recent studies have noted a hint of interaction between lexical and grammatical aspect, but again in the form of processing correlates that reflect real-time processing activities. Becker et al. (2013) found little difference between atelic verbs with the imperfective aspect and the perfective aspect in terms of the availability of events in later sentences, while telic verbs with the imperfective aspect did have an impact on the availability of events. Becker et al. claimed that with activity (atelic) verbs, the event is more homogenous throughout, while with accomplishment (telic) verbs, the event has different stages, thus event representations for imperfective and perfective are more similar for activities and less similar for accomplishments. My interpretation of Becker et al.’s results is that atelic events are a “mismatch” to the perfective aspect, because they do not have an intrinsic end-point to match the completedness of the perfective; hence, atelic activities might be considered as ongoing even with the perfective aspect. Becker et
al.’s finding of differences in effects of imperfective vs. perfective aspect with atelic verbs fits my interpretation, but differs from the findings of my study (Experiment 1). What is the difference between their study and the present study? Both tested English (so word order plays no role), but Becker et al. recorded electroencephalograms (EEG) (processing activity) while I collected participants’ preferences in the subject nouns they selected for their own sentence continuations (i.e., Source vs. Goal preference in a sentence such as John kicked the ball to Mary. He/She -----). The former records an activity reflecting a psychological reality, while the latter show readers’ explicit preferences for the focus of events after the critical aspect sentence. The brain activities may reflect an interaction that does not necessarily have a large impact on sentence interpretation. Likewise, Yap et al.’s (2009) argument for facilitation between the matching prototypical combination of lexical and grammatical aspect in sentence processing was again based on processing correlates such as reading times, rather than on any influence or change in the interpretation. Thus, the subtle interaction between lexical and grammatical aspect might be missed in referential preference (which is rather explicit), as in the present study, while processing activities (which are rather implicit) were recorded by Becker et al.

The present findings of null effects suggest a need for further investigation of the processing effects of grammatical aspect. Empirical studies that have explored grammatical aspect are currently fairly limited in number and show an imbalance in the languages they cover. English has been examined most frequently (see references in Chapter 3), but in the mainstream of psycholinguistics, even the research on this topic in English is more limited than the research that addresses other linguistic phenomena, such
as ambiguity resolution, sentence complexity, relative clauses, and so on. To illustrate this point, while processing costs are one of the major dependent measures tested in psycholinguistic studies, only Coll-Florit and Gennari’s (2001) study on Spanish and no study on English has considered processing costs and grammatical aspect until the present research (Experiment 2). Studies that target Japanese are even more rare.

Naturally, the kinds of tasks employed for testing effects of grammatical aspect for each test language have been even more limited. As discussed above, some measurements may find more effects than others, which may be related to whether they reflect subtle or deep brain activities in comprehenders. Task effects that require more attention from comprehenders during sentence processing may contribute to more effects in experiments, if comprehenders have “good enough” representations in sentence comprehension and so, as discussed, do not immediately need to commit fully to any interpretation during normal reading.

The effects of aspect are different in different languages, as the present study empirically demonstrated by showing the null effects in Japanese in contrast to the effects found in other languages, such as English. In the field of language acquisition, while most studies have attested that young children show a fixed preference for certain combinations of verbs and tense-aspect morphology, Shirai (1998) reported that Japanese children do not necessarily show such a preference. Psychological realities likewise differ from language to language and are not universally identical. This study found null effects of grammatical aspect in sentence processing in Japanese, which can be explained by the lack of opportunity for grammatical aspect to affect processing activities when it appears at the end of the sentence in incremental sentence processing. Word order provides some
explanation for the attested null effects in Japanese, but is clearly not the only factor that might affect the influence of grammatical aspect. To confirm my new claim, it is important to continue this line of research by measuring online processing activities with other verb-final languages such as Korean. As another suggestion for future research, I point out that the effect of interaction between lexical and grammatical aspect might be more powerful in child language. In other words, aspectual coercion may be absent in child language. Research on early languages has shown children to have almost universal patterns of combinations of tense-aspect morphology and lexical aspect (with the exception of Japanese; Shirai, 1998). Children begin assigning tense-aspect morphology to prototypical verbs and then extend the morphology to peripheral verbs. My Experiment 1 (English) used what we call a “peripheral” verb type, atelic events, for the target grammatical aspect, the perfective. As children are sensitive to preferred (and dispreferred) aspect cues, we may not see aspectual coercion with children; that is, perfective grammatical aspect may not function to induce children to interpret events as completed with atelic verbs.

7.3 Conclusion
I conclude that the importance of the present dissertation research is that it adds to our understanding of how we interpret language about time in events, by showing that (i) although some recent sentence processing studies have started to reveal an interaction between grammatical aspect and other sources of temporal properties (such as lexical aspect, duration of time), such an interaction has a smaller influence than grammatical aspect in terms of sentence interpretation in adult language; (ii) the psychological effects
of grammatical aspect and/or its interaction with other sources of temporal properties are affected by the relative timing of temporal cues and the construction of (other aspects of) event models; (iii) in Japanese, grammatical aspect does not have much impact on processing behaviors because the aspectual cues are available only sentence-finally, which is in line with attested patterns of early child language of Japanese, where the tense-aspect morphology is not consistently matched up with prototypical verb types, unlike in other child languages. At the same time, I proposed that the effect of grammatical aspect remains to be confirmed in a wider range of test languages and tasks, before we draw conclusions on the true effects of grammatical aspect and their interaction with other temporal sources.
## APPENDIX A

### EXPERIMENT 1, CRITICAL PREAMBLES AND PROMPTS IN TWO-ALTERNATIVE FORCED CHOICE (2AFC)

<table>
<thead>
<tr>
<th>Critical Preambles</th>
<th>Prompts in 2AFC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Experiment A (Fixed order)</td>
</tr>
<tr>
<td>1 Elizabeth {was taking/took}{a meal/meals} to Frank.</td>
<td>Elizabeth/ Frank</td>
</tr>
<tr>
<td>2 Chelsea {was passing/passed}{a sandwich/sandwiches} to Ben.</td>
<td>Chelsea/ Ben</td>
</tr>
<tr>
<td>3 Caroline {was pitching/pitched}{a ball/balls} to Jeremy.</td>
<td>Caroline/ Jeremy</td>
</tr>
<tr>
<td>4 Jerry {was returning/returned}{a sweater/sweaters} to Christine.</td>
<td>Jerry/Christine</td>
</tr>
<tr>
<td>5 David {was sending/sent}{a love/love letters} to Gina.</td>
<td>David/Gina</td>
</tr>
<tr>
<td>6 Paul {was carrying/carried}{a tray/trays} to Julia.</td>
<td>Paul/Julia</td>
</tr>
<tr>
<td>7 Claire {was hurrying/hurried}{a transcript/transcripts} to Jeff.</td>
<td>Claire/Jeff</td>
</tr>
<tr>
<td>8 Crystal {was batting/batted}{a grounder/grounders} to Jim.</td>
<td>Crystal/Jim</td>
</tr>
<tr>
<td>9 Beth {was whacking/whacked}{a wiffleball/wiffleballs} to Dennis.</td>
<td>Beth/Dennis</td>
</tr>
<tr>
<td>10 Adam {was giving/gave}{a sweater/sweaters} to Laura.</td>
<td>Adam/Laura</td>
</tr>
<tr>
<td>11 Natalie {was chucking/chucked}{a measuring tape/measuring tapes} to Jacob.</td>
<td>Natalie/Jacob</td>
</tr>
<tr>
<td>12 Tom {was bringing/brought}{a cup/cups} of tea to Stephanie.</td>
<td>Tom/ Stephanie</td>
</tr>
<tr>
<td></td>
<td>Subject</td>
</tr>
<tr>
<td>---</td>
<td>---------</td>
</tr>
<tr>
<td>13</td>
<td>Henry</td>
</tr>
<tr>
<td>14</td>
<td>Evan</td>
</tr>
<tr>
<td>15</td>
<td>Melanie</td>
</tr>
<tr>
<td>16</td>
<td>Luke</td>
</tr>
<tr>
<td>17</td>
<td>Diane</td>
</tr>
<tr>
<td>18</td>
<td>Bill</td>
</tr>
<tr>
<td>19</td>
<td>Brian</td>
</tr>
<tr>
<td>20</td>
<td>Janet</td>
</tr>
<tr>
<td>21</td>
<td>Nick</td>
</tr>
<tr>
<td>22</td>
<td>Rebecca</td>
</tr>
<tr>
<td>23</td>
<td>Angela</td>
</tr>
<tr>
<td>24</td>
<td>Josh</td>
</tr>
<tr>
<td>25</td>
<td>Mike</td>
</tr>
<tr>
<td>26</td>
<td>Kyle</td>
</tr>
<tr>
<td>27</td>
<td>Linda</td>
</tr>
<tr>
<td>28</td>
<td>Molly</td>
</tr>
<tr>
<td>29</td>
<td>Karen</td>
</tr>
<tr>
<td>Line</td>
<td>Action</td>
</tr>
<tr>
<td>------</td>
<td>--------</td>
</tr>
<tr>
<td>30</td>
<td>Anne</td>
</tr>
<tr>
<td>31</td>
<td>Kim</td>
</tr>
<tr>
<td>32</td>
<td>Tim</td>
</tr>
<tr>
<td>33</td>
<td>Sue</td>
</tr>
<tr>
<td>34</td>
<td>Brandon</td>
</tr>
<tr>
<td>35</td>
<td>Rachel</td>
</tr>
<tr>
<td>36</td>
<td>Jennifer</td>
</tr>
<tr>
<td>37</td>
<td>Kara</td>
</tr>
<tr>
<td>38</td>
<td>Richard</td>
</tr>
<tr>
<td>39</td>
<td>Keith</td>
</tr>
<tr>
<td>40</td>
<td>Ethan</td>
</tr>
<tr>
<td>41</td>
<td>Bruce</td>
</tr>
<tr>
<td>42</td>
<td>Naomi</td>
</tr>
<tr>
<td>43</td>
<td>Seth</td>
</tr>
<tr>
<td>44</td>
<td>Eric</td>
</tr>
<tr>
<td>45</td>
<td>Andrew</td>
</tr>
<tr>
<td>46</td>
<td>Emily</td>
</tr>
<tr>
<td></td>
<td>Subject (Past Tense)</td>
</tr>
<tr>
<td>---</td>
<td>---------------------</td>
</tr>
<tr>
<td>47</td>
<td>Alan was towing/towed an old jalopy/old jalopies</td>
</tr>
<tr>
<td>48</td>
<td>Larry was heaving/heaved a box/boxes</td>
</tr>
<tr>
<td>49</td>
<td>Maria was carting/carted an old computer/old computers</td>
</tr>
<tr>
<td>50</td>
<td>John was donating/donated a toy/toys</td>
</tr>
<tr>
<td>51</td>
<td>Joe was slugging/slugged a line drive/line drives</td>
</tr>
<tr>
<td>52</td>
<td>Kelly was raising/raised a ladder/ladders to Daniel</td>
</tr>
<tr>
<td>53</td>
<td>Chad was feeding/fed a piece/pieces of chocolate to Nancy</td>
</tr>
<tr>
<td>54</td>
<td>Brad was throwing/threw a hat/hats to Margaret</td>
</tr>
<tr>
<td>55</td>
<td>Sophia was lowering/lowered a first aid kit/first aid kits to Joel</td>
</tr>
<tr>
<td>56</td>
<td>Jason was nudging/nudged a microphone/microphones to Hannah</td>
</tr>
<tr>
<td>57</td>
<td>George was slapping/slapped a beach/beach balls to Sarah</td>
</tr>
<tr>
<td>58</td>
<td>Kevin was delivering/delivered a letter/letters to Shelly</td>
</tr>
<tr>
<td>59</td>
<td>Caitlin was lobbing/lobbed a football/footballs to Mitch</td>
</tr>
<tr>
<td>60</td>
<td>Albert was flicking/flicked a paperclip/paperclips to Kristen</td>
</tr>
</tbody>
</table>
APPENDIX B

SUB-ANALYSES IN EXPERIMENT 1

Iterative Items. Figure B.1 shows the mean Source ratio across conditions by iterative items and by less-iterative items according to the norming study (See Section 4.1.2). Paired t-tests revealed that in the 32 most iterative items the imperfective condition increased Source bias (by participant, 27.60%) more than in either of the perfective conditions with the singular noun (by participant, 21.12%) both by participants and by items ($t_1(71) = 3.03, p = 0.0033; t_2(31) = 3.20, p = 0.0031$, two-tailed), or with the plural noun (by participant, 21.99%) both by participants and by items ($t_1(71) = 2.60, p = 0.011; t_2(31) = 2.86, p = 0.0076$, two-tailed). Non-iterative items revealed no statistical differences in Source preference across the three conditions.

Paired t-tests for comparison of Source ratio between iterative versus non-iterative items on subject analyses in each condition revealed that the non-iterative items in the plural-perfective condition increased Source bias, 26.91%, more than iterative items, 21.99%, $t_1(71) = 2.78, p = 0.0069$, two-tailed. Similarly, in the perfective-singular condition, non-iterative items showed marginal increase of Source bias, 25.41% (by participant), over iterative items, 21.12% (by participant), $t_1(71) = 1.95, p = 0.055$, two-tailed. No difference was made by iterativeness in the imperfective condition ($t_1(71) = 0.32, p = n.s.$).
Figure B.1. Mean source preference ratio by participants after reading a Perfective-Singular, Perfective-Plural, or Imperfective-Singular preamble, and 95% confidence intervals with distinction of iterativeness in Experiment 1.

**Co-located Items.** Figure B.2 shows the mean Source ratio across conditions by co-located items and by less-co-located items according to the norming study (See Section 4.1.2). Paired t-tests between conditions in the 42 most co-located items showed that the imperfective-singular condition increased Source bias (by participant, 26.54%) over the perfective-singular condition (by participant, 20.99%) both by participants and items ($t_1(71) = 2.69, p=0.0089$; $t_2(41) = 3.29, p=0.0021$, two-tailed). There was a suggestion that the imperfective-singular condition increased the Source bias more than the perfective-plural (by participant, 23%) ($r^2: p=0.086$), and the perfective-plural increased the Source bias more than the perfective-singular ($r^2: p=0.082$), but both were only by items. The 18 less-co-located items revealed no differences in Source bias across the three conditions, except for a suggestion of increase of Source bias in the imperfective-
singular (by item, 29.86%) more than the perfective-plural (by item, 24.77%) only by items ($t_{2}(17) = 2.50, p=0.023$, two-tailed).

Paired $t$-tests for comparison of Source bias between co-located items and less-co-located items on subject analyses in each condition revealed that the less-co-located items in the singular conditions increased Source bias more than the co-located items (with the perfective, 29.26% vs. 20.99%; $t_{1}(71)=2.79, p=0.0068$, two-tailed; with the imperfective, 33.71% vs. 26.54%; $t_{1}(71)=2.29, p=0.025$, two-tailed). Interestingly, the plural-perfective showed no difference (23.24% vs. 27.11%; $t_{1}(71) = 1.22, p=\text{n.s.}$).

**Figure B.2.** Mean source preference ratio by participants after reading a Perfective-Singular, Perfective-Plural, or Imperfective-Singular preamble, and 95% confidence intervals with distinction of co-locatedness in Experiment 1.

**Pronoun vs. Noun.** Figure B.3 shows the mean Source ratio across conditions by the 24 pronoun items and the 36 “proper name”\(^{45}\) items from Experiments A and B combined.

\(^{45}\) As mentioned in Section 4.1.2, note that these 36 items were used with proper name referents in Experiment A, while the same items were used with pronoun referents in Experiment B (as Experiment B was an all-pronoun referents experiment).
Paired \( t \)-tests between conditions in each item group showed that the imperfective condition increased the Source bias over the perfective condition, in both the perfective-singular condition in the pronoun items (29.34\% vs. 22.57 \% (by participant), \( t^1(71) = 2.89: p = 0.0051, t^2(23) = 2.73: p = 0.012 \), two-tailed) and the perfective-plural condition in the proper name items (27.66\% vs. 23.38 \% (by participant), \( t^1(71) = 2.25: p = 0.027, t^2(35) = 2.11: p = 0.042 \), two-tailed). The results showed no difference in Source bias between singular and plural nouns within the perfective.

Paired \( t \)-tests for comparison of Source bias between pronoun items and proper name items on subject analyses in each condition revealed no differences in either condition.

![Source preference ratio](image)

**Figure B.3.** Mean source preference ratio by participants after reading a Perfective-Singular, Perfective-Plural, or Imperfective-Singular preamble, and 95\% confidence intervals with distinction of Pronoun set vs. Noun set in Experiment 1
## APPENDIX C

### EXPERIMENT 2 (ENGLISH), CRITICAL SENTENCES

<table>
<thead>
<tr>
<th>Sentence</th>
<th>Verb Phrase</th>
<th>Quantity</th>
<th>Activity</th>
<th>Location/Condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Larissa/ was listening to/listened to/</td>
<td>{four songs/24 songs}/ on shuffle/</td>
<td>during her morning walk.</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Henry/ was sprinting/sprinted/</td>
<td>{100 yards/400 yards}/ in the youth competition/</td>
<td>the year/ he/ was awarded/ a scholarship.</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Luke/ was skating/skated/</td>
<td>{500 m/10,000 m}/ at full speed/</td>
<td>when/ he/ won/ the race.</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Mr. Williams/ was writing/wrote/</td>
<td>{two checks/15 checks}/ in his office/</td>
<td>yesterday/ when/ his aunt/ called.</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Patrick/ was climbing/climbed/</td>
<td>{two flights/ten flights}/ of stairs/</td>
<td>as a warm up exercise/ early / this morning.</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Sean/ was doing/did/</td>
<td>{five push-ups/30 push-ups}/ as a part/</td>
<td>of his routine/ before/ he/ went/ to bed.</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Tom/ was golfing/golfed/</td>
<td>{three holes/18 holes}/ with his colleagues/</td>
<td>last Sunday.</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Audrey/ was burning/burned/</td>
<td>{a hole/16 holes}/ in her favorite rug/</td>
<td>yesterday.</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Chelsea/ was eating/ate/</td>
<td>{two strawberries/12 strawberries}/</td>
<td>quickly/ after class.</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Naomi/ was watching/watched/</td>
<td>{four clips/20 clips}/ on YouTube/</td>
<td>silently/ with headphones/ in her room.</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Samantha/ was vacuuming/vacuumed/</td>
<td>{a room/six rooms}/ in the dormitory/</td>
<td>yesterday.</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Stephanie/ was drawing/drew/</td>
<td>{four flowers/40 flowers}/ on the blackboard/</td>
<td>when/ the teacher/ walked in.</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>Conrad/ was screaming/screamed/</td>
<td>{for three minutes/for 30 minutes}/</td>
<td>at Gayle/ to leave him/ alone.</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Ethan/ was sneezing/sneezed/</td>
<td>{for three minutes/for 30 minutes}/</td>
<td>because of hay fever/ while/ doing / his homework.</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>Frank/ was bathing/bathed/</td>
<td>{for ten minutes/for 60 minutes}/ in a herbal bath/</td>
<td>to relax.</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>Isaac/ {was sleeping/slept}/ {for two hours/for 12 hours}/ in his room/ yesterday.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>Jeremy/ {was dancing/danced}/ {for an hour/for four hours}/ at a dance club/ near the beach.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>Brian/ {was waiting/waited}/ {for three days/for 14 days}/ for her reply/ in excitement.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>Fiona/ {was laughing/laughed}/ {for a minute/for 20 minutes}/ to herself/ this afternoon.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>Irene/ {was working/worked}/ {for two weeks/for 20 weeks}/ in the lab/ by herself.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>Joanne/ {was swimming/swam}/ {for five minutes/for 30 minutes}/ in a crowded pool/ enthusiastically.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>Josephine/ {was hiking/hiked}/ {for two hours/for seven hours}/ in the Rockies/ in fine weather.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>Karen/ {was typing/typed}/ {for an hour/for five hours}/ to meet/ her deadline.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>Rebecca/ {was singing/sang}/ {for five minutes/for 40 minutes}/ in the shower/ yesterday evening.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
# EXPERIMENT 3 (JAPANESE, PROGRESSIVE), CRITICAL SENTENCES

<table>
<thead>
<tr>
<th>No.</th>
<th>Text (Japanese)</th>
<th>Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>竜奈は/散歩中に/{四曲の歌を/24曲の歌を}うっとりと/{聞いていた。/聞いた。}</td>
<td>Rena {was listening/listened} to {four/24} songs absorbedly during a walk.</td>
</tr>
<tr>
<td>2</td>
<td>一輝は/校庭で/{50mを/800mを}真剣に/{走っていた。/走った。}</td>
<td>Kazuki {was running/ran} {50m/800m} seriously in the schoolyard.</td>
</tr>
<tr>
<td>3</td>
<td>健一は/スピードスケートで/{500メートル/10000メートルを}全力で/{滑っていた。/滑った。}</td>
<td>Kenichi {was skating/skated} {500 m/10,000 m} at full speed.</td>
</tr>
<tr>
<td>4</td>
<td>和彦は/オフィスで/{二通の証明書を/15通の証明書を}事務的に/{書いていた。/書いた。}</td>
<td>Kazuhiko {was writing/wrote} {two/15} certificates in his office in a businesslike manner.</td>
</tr>
<tr>
<td>5</td>
<td>学は/その休暇中に/{二冊の本を/10冊の本を}もくもくと/{読んでいた。/読んだ。}</td>
<td>Manabu {was reading/read} {two/10} books in silence during vacation.</td>
</tr>
<tr>
<td>6</td>
<td>拓也は/図書室で/{五件の文献を/30件の文献を}丁寧に/{調べていた。/調べた。}</td>
<td>Takuya {was consulting/consulted} {five/30} copies of literature carefully in the library.</td>
</tr>
<tr>
<td>7</td>
<td>雄太は/アルバイトで/{三台の机を/18台の机を}たくみに/{運んでいた。/運んだ。}</td>
<td>Yuta {was carrying/carried} {three/18} desks skillfully at his part-time job.</td>
</tr>
<tr>
<td>8</td>
<td>成美は/今朝/{12枚の皿を/80枚の皿を}慎重に/{数えていた。/数えた。}</td>
<td>Narumi {was counting/counted} {12/80} dishes carefully this morning.</td>
</tr>
<tr>
<td>9</td>
<td>沙織は/ひな祭りに/{二杯の甘酒を/10杯の甘酒を}さりげなく/{飲んでいた。/飲んだ。}</td>
<td>Saori {was drinking/drank} {two/10} cups of Amazake casually at the Girl’s Festival.</td>
</tr>
<tr>
<td>10</td>
<td>八重子は/ユーチューブで/{四個の動画を/20個の動画を}楽しく/{見ていた。/見た。}</td>
<td>Yaeko {was watching/watched} {four/20} video clips on YouTube enjoyably yesterday.</td>
</tr>
<tr>
<td>No.</td>
<td>Japanese Text</td>
<td>English Translation</td>
</tr>
<tr>
<td>-----</td>
<td>---------------</td>
<td>---------------------</td>
</tr>
<tr>
<td>11</td>
<td>智美は昨日六部屋の客室をきれいに掃いていた。</td>
<td>‘Satomi was sweeping/swept {six/20} visitor’s rooms clean yesterday.’</td>
</tr>
<tr>
<td>12</td>
<td>裕子は黒板に四個の星を上手に描いていた。</td>
<td>‘Yuko was drawing/drew {four/40} stars well on the blackboard.’</td>
</tr>
<tr>
<td>13</td>
<td>声は電話で三分間明日の予定を話していた。</td>
<td>‘Ryo was talking/talked for {three/30} minutes about tomorrow’s plan on the phone.’</td>
</tr>
<tr>
<td>14</td>
<td>哲也は窓の近くで五分間静かにすすんでいた。</td>
<td>‘Tetsuya was cooling/cooled off for {five/40} minutes quietly by the window.’</td>
</tr>
<tr>
<td>15</td>
<td>達也は薬湯に十分間浸かっていた。</td>
<td>‘Tatsuya was soaking/soaked for {ten/60} minutes deeply in a herbal bath.’</td>
</tr>
<tr>
<td>16</td>
<td>淳は昨日彼の部屋で眠っていた。</td>
<td>‘Atsushi was sleeping/slept for {two/12} hours in his room yesterday.’</td>
</tr>
<tr>
<td>17</td>
<td>康平は週末に一時間/四時間はつらつと踊っていた。</td>
<td>‘Kohei was dancing/danced {one/four} hour(s) actively on the weekend.’</td>
</tr>
<tr>
<td>18</td>
<td>剛は専務の返事を三日間おおらかに待っていた。</td>
<td>‘Tsuyoshi was waiting/waited for {three/14} days for a reply from the Executive Director with patience.’</td>
</tr>
<tr>
<td>19</td>
<td>久美子は今日の午後一分間ひっそりと笑っていた。</td>
<td>‘Kumiko was laughing/laughed for {one/20} minutes silently this afternoon.’</td>
</tr>
<tr>
<td>20</td>
<td>明美はその研究室で二週間/二週間独りきりで働くていた。</td>
<td>‘Akemi was working/worked for {two/20} weeks alone in the laboratory.’</td>
</tr>
<tr>
<td>21</td>
<td>舞は大学のプールで五分間/三十分間熱心に游泳していた。</td>
<td>‘Mai was swimming/swam for {five/30} minute keenly in the university pool.’</td>
</tr>
<tr>
<td>22</td>
<td>順子はこども部屋で二時間/七時間楽しく遊んでいた。</td>
<td>‘Junko was playing/played for {two/seven} hours in the child room happily.’</td>
</tr>
<tr>
<td>23</td>
<td>麻衣は友人と一時間/五時間草原をぶらついていた。</td>
<td>‘Mai was wandering/wandered for {one/five} hour(s) in the grassland with her friends.’</td>
</tr>
<tr>
<td>24</td>
<td>佳奈はお昼に五分間/四十分間台所で歌っていた。</td>
<td>‘Kana was singing/sang for {five/40} minutes in the kitchen at lunch time.’</td>
</tr>
</tbody>
</table>
APPENDIX E

EXPERIMENT 4 (JAPANESE, RESULTATIVE), CRITICAL SENTENCES

<table>
<thead>
<tr>
<th>No.</th>
<th>Japanese Description</th>
<th>English Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>コピー機は不便にも三日間/14日間/すっかり/壊れていた。壊れた。</td>
<td>‘To his dismay, the copier {had been/was} completely broken for {three/14} days.’</td>
</tr>
<tr>
<td>2</td>
<td>哲也はその同好会で三ヶ月の間/12ヶ月の間/会長に決まった。</td>
<td>‘Tetsuya {had been/was} selected for {three/12} months as the chairman at the club.’</td>
</tr>
<tr>
<td>3</td>
<td>タンスの跡は畳に三分間/50分間/しっかりと残っていた。</td>
<td>‘The mark of the chest {had remained/remaining} for {three/50} minutes clearly on the tatami mat.’</td>
</tr>
<tr>
<td>4</td>
<td>久美子は同僚と二ヶ月間/15ヶ月間/ひそかに婚約していた。</td>
<td>‘Kumiko {had been/was} engaged for {two/15} months secretly to her colleague.’</td>
</tr>
<tr>
<td>5</td>
<td>健太は友人と一時間/五時間/花見に出かけていた。</td>
<td>‘Kenta {had been/went} out for {one/five} hour(s) to see the cherry blossoms with his friends.’</td>
</tr>
<tr>
<td>6</td>
<td>恵はその夏一週間ほど/八週間ほど/故郷に戻っていた。</td>
<td>‘Megumi {had been/was} back for about {one/eight} week(s) that summer to her home town.’</td>
</tr>
<tr>
<td>7</td>
<td>美穂はカバンに三日間ほど/10日間ほど/バッジをつけていた。</td>
<td>‘Miho {had worn/wore} a badge on her bag for about {three/10} days.’</td>
</tr>
<tr>
<td>8</td>
<td>その作品は美術館に一週間/12週間/晴れればと並んでいた。</td>
<td>‘The art work {had been/was} shown in the museum for {one/12} week(s) proudly.’</td>
</tr>
<tr>
<td>9</td>
<td>和彦はアパートを一週間ほど/七週間ほど/すっかり空けていた。</td>
<td>‘Kazuhiko {had emptied/emptied} his apartment for {one/seven} week(s) completely.’</td>
</tr>
<tr>
<td>10</td>
<td>一輝は壇上に五分間/40分間/悠然と立っていた。</td>
<td>‘Kazuki {had stood/stood} on a stage for {five/40} minutes sedately.’</td>
</tr>
<tr>
<td>11</td>
<td>健一は仕事を二ヶ月間/八ヶ月間/病気で辞めていた。</td>
<td>‘Kenichi {had quit/quit} his job for {two/eight} months on account of sickness.’</td>
</tr>
<tr>
<td>行目</td>
<td>文章</td>
<td></td>
</tr>
<tr>
<td>------</td>
<td>-------</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>明美は / スケートリンクの端に / (三分ほど/20分ほど) / 静かに / {止まっていた。/止まった。} / 'Akemi {had stopped/stopped} at the end of the skating rink for about {three/20} minutes quietly.'</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>患者は / ベッドに / (五分ほど/30分ほど) / にこやかに / {起き上がっていた。/起き上がった。} / 'The patient {had got/got} up for about {five/30} minutes on the bed with a smile.'</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>大輪の花は / ベランダに / (二日ほど/七日ほど) / 見事に / {咲いていた。/咲いた。} / 'The large flower {had bloomed/bloomed} for about {two/seven} days stunningly at the veranda.'</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>小さな灯りは / ソファの近くに / (五分間/40分間) / ほんのりと / {ともっていた。/ともった。} / 'The small light {had been/was} lighted for {five/40} minutes near the sofa faintly.'</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>虹は / 向こうの空に / (二分間ほど/15分間ほど) / うっすらと / {出ていた。/出た。} / 'Rainbow {had appeared/appeared} for about {two/15} minutes in the sky dimly.'</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>達也は / いとこに / (三日間/14日間) / スーツケースを / {貸していた。/貸した。} / 'Tatsuya {had lent/lent} his suitcase for {three/14} days to his cousin.'</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>霧は / 山頂付近で / (10分間/50分間) / すっかり / {消えていた。/消えた。} / 'The fog {had been/was} disappeared for {10/50} minutes completely near the summit.'</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>電話は / どうしてか / (一分間/30分間) / 非常に / {聞こえづらくなっていました。/聞こえづらくなった。} / 'It {had become/became} very difficult to hear on the phone for {one/30} minute(s) for some reason.'</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>雨は / その日 / (二時間/20時間) / ばったりと / {やんでいた。/やんだ。} / 'The rain {had ceased/ceased} for {two/20} hours completely on the day.'</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>西の空は / 茜色に / (五分間ほど/30分間ほど) / きれいに / {染まっていました。/染まった。} / 'The west sky {had dyed/dyed} in madder red for about {five/30} minutes beautifully.'</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>金庫の鍵は / この前 / (二日ほど/20日ほど) / ひそかに / {なくなっていた。/なかった。} / 'The key to the safe {had gone/went} missing for about {two/20} days covertly recently.'</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>老人は / 通路側の席に / (一時間ほど/五時間ほど) / ゆっくりと / {座っていました。/座った。} / 'An old man {had sat/sat} on the aisle seat for about {one/five} hour(s) in easy state.'</td>
<td></td>
</tr>
</tbody>
</table>
| 24 | 八重子は着物の裾を（三秒ほど/30秒ほど）うっかり（踏んでいた。/踏んだ。）
|    | ‘Yaeko {had stepped/stepped} on the hem of kimono for about {three/30} seconds carelessly.’ |
APPENDIX F

EXPERIMENT 5 (JAPANESE, CONTROL), CRITICAL SENTENCES

The first 12 sentences are resultative sentences and the latter 12 sentences are progressive sentences.

<table>
<thead>
<tr>
<th>Sentence</th>
<th>Resultative Sentence</th>
<th>Progressive Sentence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>コピー機は/不便にも/すっかり/{壊れていた。/壊れた。}</td>
<td>To his dismay, the copier {had been/was} completely broken.</td>
</tr>
<tr>
<td>2</td>
<td>タンスの跡は/畳に/しっかりと/{残っていた。/残った。}</td>
<td>The mark of the chest {had remained/remained} clearly on the tatami mat.</td>
</tr>
<tr>
<td>3</td>
<td>恵は/その夏/故郷に/{戻っていた。/戻った。}</td>
<td>Megumi {had been/was} back that summer to her home town.</td>
</tr>
<tr>
<td>4</td>
<td>その作品は/美術館に/晴ればれと/{並んでいた。/並んだ。}</td>
<td>The art work {had been/was} shown in the museum proudly.</td>
</tr>
<tr>
<td>5</td>
<td>一輝は/壇上に/悠然と/{立っていた。/立った。}</td>
<td>Kazuki {had stood/stood} on a stage sedately.</td>
</tr>
<tr>
<td>6</td>
<td>大輪の花は/ベランダに/見事に/{咲いていた。/咲いた。}</td>
<td>The large flower {had bloomed/bloomed} stunningly at the veranda.</td>
</tr>
<tr>
<td>7</td>
<td>小さな灯りは/ソファの近くに/ほんのりと/{ともっていた。/ともった。}</td>
<td>The small light {had been/was} lighted on near the sofa faintly.</td>
</tr>
<tr>
<td>8</td>
<td>虹は/向こうの空に/うっすらと/{出ていた。/出た。}</td>
<td>Rainbow {had appeared/appeared} in the sky dimly.</td>
</tr>
<tr>
<td>9</td>
<td>霧は/山頂付近で/すっかり/{消えていた。/消えた。}</td>
<td>The fog {had been/was} disappeared completely near the summit.</td>
</tr>
<tr>
<td>10</td>
<td>雨は/その日/ぱたりと/{やんでいた。/やんだ。}</td>
<td>The rain {had ceased/ceased} completely on the day.</td>
</tr>
<tr>
<td>11</td>
<td>西の空は/茜色に/きれいに/{染まってい。/染まった。}</td>
<td>The west sky {had dyed/dyed} in madder red beautifully.</td>
</tr>
<tr>
<td>12</td>
<td>金庫の鍵は/この前/ひそかに/{なくなっていた。/なくなった。}</td>
<td>The key to the safe {had gone/went} missing covertly recently.</td>
</tr>
<tr>
<td>13</td>
<td>亮は/電話で/友人と/{話していた。/話した。}</td>
<td>Ryo {was talking/talked} with his friend on the phone.</td>
</tr>
<tr>
<td>14</td>
<td>哲也は/窓の近くで/静かに/{すずんでいた。/すずんだ。}</td>
<td>Tetsuya {was cooling/cooled} off quietly by the window.</td>
</tr>
<tr>
<td>15</td>
<td>達也は薬湯にしみじみと浸かっていた。浸かった。</td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>淳は昨日彼の部屋で眠っていた。眠った。</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>康平は週末にはつらつと踊っていた。踊った。</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>剛は専務の返事をおおらかに待っていた。待った。</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>久美子は今日の午後ひっそりと笑っていた。笑った。</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>明美はその研究室で独りきりで働いていた。働いた。</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>舞は大学のプールで熱心に泳いでいた。泳いだ。</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>順子はこども部屋で楽しく遊んでいた。遊んだ。</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>麻衣は友人と草原をぶらついていた。ぶらついた。</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>佳奈はお昼に台所で歌っていた。歌った。</td>
<td></td>
</tr>
</tbody>
</table>

'Tatsuya {was soaking/soaked} deeply in a herbal bath.'
'Atsushi {was sleeping/slept} in his room yesterday.'
'Kohei {was dancing/danced} actively on the weekend.'
'Tsuyoshi {was waiting/waited} for a reply from the Executive Director with patience.'
'Kumiko {was laughing/laughed} silently this afternoon.'
'Akemi {was working/worked} alone in the laboratory.'
'Mai {was swimming/swam} keenly in the university pool.'
'Junko {was playing/played} in the child room happily.'
'Mai {was wandering/wandered} in the grassland with her friends.'
'Kana {was singing/sang} in the kitchen at lunch time.'
**APPENDIX G**

**EXPERIMENTS 6 AND 8 (PROGRESSIVE), CRITICAL SENTENCES**

|  | 玲奈はバイオリンが好きだ。 | Rena likes violins.  
玲奈は先週、コンサートに出かけた。 | Last week, Rena went to the concert.  
玲奈は音色の良さに感嘆した。 | Rena admired the quality of the sound.  
玲奈はうっとりと演奏を聴いていた。 | ‘Rena {was listening/listened} to the playing absorbedly.’  
玲奈はうっとりと演奏を聴いた。 |  
選曲はバッハとシューベルトだった。 | Music of Bach and Schubert were played.  
バイオリンは音の表情が豊かだった。 | The sound of the violin was rich.  
玲奈は元々ピアノも好きだった。 | Rena also liked the piano.  
バイオリンはピアノを伴奏にしていた。 | The violin accompanied the piano performance.  
玲奈は二種類の楽器が楽しめる点が気に入った。 | Rena liked it as she could enjoy the two types of instruments.  
玲奈は CD を買って帰った。 | Rena bought a CD and went back home.  
剛は郊外に住んでいた。 | Tsuyoshi was living in the suburbs.  
近々、近所でテレビ番組の公開録画が行われることになった。 | The public recording of a TV program was to be carried out soon in the neighborhood.  
剛は観覧に応募した。 | Tsuyoshi applied for a viewing.  
剛は期待して通知を待っていた。 | ‘Tsuyoshi {was waiting/waited} for the notification with eagerness.’  
二週間後にはがきが届くらしい。 | A postcard notification is expected to arrive in two weeks.  
公開録画には昼の部と夜の部があっ | There were a matinee and a soiree for the public recording.  
た。 | (He was) intrigued with the matinee.  
昼の部のほうが興味をそそった。 | Tsuyoshi had signed up for the matinee.  
剛は昼の部に申し込んでいた。 | Tsuyoshi won it with flying colors.  
剛は見事当選した。 | On the day, Tsuyoshi was video-taped in the audience seats.  
剛は当日、観客席の中に写ることになった。 |  
健一はスキーの中級者だった。 | Kenichi was an intermediate-level skier.  
健一の伯父といとは一級の腕前だった。 | Kenichi’s uncle and cousin were first-
<p>| 先日には親戚一同でスキーを楽しんだ。 | rate skiers. (He) enjoyed skiing with all his relatives the other day. |
| 健一は上級者コースを（滑っていた。/滑った。）&lt;br&gt;‘Kenichi {was skiing/skied} on the advanced course.’ | 健一には絶壁のように見えた。&lt;br&gt;ふもとまで下るのは大変だった。&lt;br&gt;伯父といとこは颯爽と滑っていた。&lt;br&gt;健一は本格的にスキーを習うことにした。 |
| それは急斜面だった。&lt;br&gt;健一には絶壁のように見えた。&lt;br&gt;ふもとまで下るのは大変だった。&lt;br&gt;伯父といとこは颯爽と滑っていた。&lt;br&gt;健一は本格的にスキーを習うことにした。 | It was a steep slope.&lt;br&gt;It looked like a cliff to Kenichi.&lt;br&gt;It was hard to ski down to the bottom of the hill.&lt;br&gt;(His) uncle and cousin had skied beautifully.&lt;br&gt;Kenichi decided to learn to ski in earnest. |
| 直美は翔と知り合いになった。&lt;br&gt;直美は翔とごくたまにメールをやりとりする。&lt;br&gt;知り合いになって一ヶ月が経った。 | Naomi became acquainted with Sho.&lt;br&gt;Naomi and Sho exchange e-mails rarely.&lt;br&gt;One month has passed since (they) became acquainted. |
| 直美は久しぶりに翔にメールを（書いていた。/書いた。）&lt;br&gt;‘Naomi {was writing/wrote} a letter to Sho after a long while.’ | それは慰労会のお知らせだった。&lt;br&gt;それぞれの友人にも声を掛けあっていた。&lt;br&gt;二人とも顔が広かった。&lt;br&gt;慰労会には新しい面々も多く集まった。&lt;br&gt;こうして親睦が深まっていった。 |
| 直美は翔とごくたまにメールをやりとりする。&lt;br&gt;知り合いになって一ヶ月が経った。 | It was an announcement of a thank-you party for employees.&lt;br&gt;(They) each informed (their) friends of it.&lt;br&gt;They both knew a lot of people.&lt;br&gt;Many new faces gathered at the thank-you party.&lt;br&gt;This way friendship deepened. |
| 学は読書家だった。&lt;br&gt;学の職場は先週から休暇に入った。&lt;br&gt;学は自宅で読書三昧だった。 | Manabu was a great reader.&lt;br&gt;Manabu’s workplace went on vacation last week.&lt;br&gt;Manabu was absorbed in reading at home. |
| 学はもくもくと本を（読んでいた。/読んだ。）&lt;br&gt;‘Manabu {was reading/read} books silently.’ | うち二冊は文庫本だった。&lt;br&gt;学は特に二冊目が気に入った。&lt;br&gt;それは詩集だった。&lt;br&gt;学は詩も好きだった。&lt;br&gt;なかでも青春や哲学をテーマにしたものが好きだった。&lt;br&gt;学は続いて推理小説に没頭した。 | Among two of the books were Bunko-bon (pocket editions).&lt;br&gt;Manabu liked the second book in particular.&lt;br&gt;It was a collection of poems.&lt;br&gt;Manabu liked poetry as well.&lt;br&gt;Among others, (he) liked those with themes on youth or philosophy. |</p>
<table>
<thead>
<tr>
<th>6</th>
<th>拓也はレポートに取り組んでいた。先行研究をまとめる課題だった。拓也は昨日、図書館に足を運んだ。</th>
<th>Takuya was working on reports. It was an assignment to summarize previous research. Yesterday, Takuya went to the library.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>拓也は丁寧に文献を調べていた。</td>
<td>‘Takuya was consulting the literature carefully.’</td>
</tr>
<tr>
<td></td>
<td>その中に興味を引くテーマがあった。拓也はそのテーマを発展させたいと思った。拓也は時間惜しんで取り組んでいた。拓也のレポートの完成度は高かった。担当教授から、卒論につなげてみても提案があった。</td>
<td>There was a theme that interested him. Takuya wanted to develop that theme. Takuya was working every moment. Takuya’s report had a high completion level. His adviser suggested him to extend it to a graduation thesis.</td>
</tr>
<tr>
<td>7</td>
<td>雄太は一人暮らしていた。雄太は先月、引っ越しをした。友人二人に手伝ってもらった。</td>
<td>Yuta was living alone. Last month, Yuta moved. (He) had two friends help him.</td>
</tr>
<tr>
<td></td>
<td>雄太は荷物を小分けにしていった。友人二人とバケツリレーのようにした。意外と荷物は多かった。雄太は二人の好きなものをご馳走した。体力を使ったので三人とも食欲が旺盛だった。</td>
<td>Yuta had subdivided the luggage. It was carried with two friends in a bucket brigade manner. Unexpectedly, there was a lot of luggage. Moving finished in less than an hour thanks to them. Yuta treated them to their favorite food. They all had big appetites because they had used energy.</td>
</tr>
<tr>
<td>8</td>
<td>成美は先週、学会に参加した。初日の夕方には懇談会があった。成美は懇談会の準備を手伝った。</td>
<td>Last week, Narumi took part in an academic conference. There was a reception on the first evening. Narumi helped prepare for the reception.</td>
</tr>
<tr>
<td></td>
<td>成美は慎重に皿を数えていた。</td>
<td>‘Narumi was counting dishes carefully.’</td>
</tr>
<tr>
<td></td>
<td>係員は割り箸を準備していた。料理は予定の時間に配達されていた。</td>
<td>Attendants were preparing chopsticks. Food had been delivered at the</td>
</tr>
</tbody>
</table>
めいめいが着々と仕事をこなした。準備が済むと、成美は懇親会に加わった。会場には大勢の知り合いが居た。恩師に再会できたのは望外の喜びだった。

<table>
<thead>
<tr>
<th>第9行</th>
</tr>
</thead>
</table>
| 沙織は勉強に忙しかった。三時間、机に向かって集中していた。沙織は少し休憩をとった。

Saori was busy studying. (She) was focused at her desk for three hours. Saori took a little break.

沙織はココアを飲んでいた。/飲んだ。

‗Saori {was drinking/drank} cocoa.‘

その日は濃い目に作った。体が芯から温まってきた。リラックスするとやる気が出てくる。沙織の頭は再び冴えてきた。彼女は引き続き勉強に取り掛かった。順調に明日の分まで済ませることができた。

<table>
<thead>
<tr>
<th>第10行</th>
</tr>
</thead>
</table>
| 恵子は留学の準備をしていた。彼女は語学力を強化していた。恵子はインターネットを活用した。

Keiko was preparing for studying abroad. She was strengthening her language skills. Keiko was making use of the Internet.

恵子はニュースサイトで動画を見た。

‗Keiko {was watching/watched} videos on the news website.‘

さまざまなテレビ局がホームページを持っていた。ニュースは毎日、内容が変わることが良かった。そして最新の語法に触れることができた。音質も数年前にくらべて格段に良い。恵子は便利な世の中になったと思った。

<table>
<thead>
<tr>
<th>第11行</th>
</tr>
</thead>
</table>
| 哲也は週末は自宅で過ごす。その週末には掃除をしていた。

Tetsuya spends weekends at home. (He) was cleaning that weekend. First (he) started with the study room.
まずは勉強部屋から始めた。
哲也はせっせと床を掃いていた。/掃いた。
‘Tetsuya {was sweeping/swept} the floor hard.’

案外ほこりが溜まっていった。
哲也はすこし咳き込んでしまった。
ふと床の上に何かを発見した。
それは紛失したネジだった。
掃除機だったら吸い込んでいたところだった。
Dust had accumulated more than he had expected.
It made Tetsuya cough a little.
Tetsuya thought he should clean frequently.
Suddenly (he) found something on the floor.
It was a screw (he) had lost.
It would have been vacuumed if he had used a vacuum cleaner.

裕子は夕食の準備をしていた。
彼女の得意なシチューを煮込むことにした。
裕子は段取りよく進めた。
Yuko was preparing dinner.
(She) decided to cook a stew, one of her specialties.
Yuko efficiently proceeded in cooking.

裕子は気分良く鼻歌を歌っていた。
鍋にルーを入れようとした。
裕子はルーを切らしていた。
こういう時には経験が役に立った。
おかげで、かえってコクのあるシチューやが完成した。
(She) tried to add roux in the pan.
Yuko had run out of roux.
(Y)Our past experience became handy in a case like this.
Yuko made it from scratch with butter and flour.
Thanks to it, her stew was rather rich.

亮は旅行会社に勤めていた。
彼の提案する企画は毎回好評だった。
亮は春季の企画を練った。
Ryo worked at a travel agency.
His proposed plans were well received each time.
Ryo worked out a plan for spring.

亮は新企画について取引先と話していた。
それは桜の名所を訪れるものだった。
外国旅行者がターゲットだった。
話し合いは大いに盛り上がった。
企画は実現することとなった。
It was to visit places famous for cherry blossoms.
The novelty was that English-speaking guides would accompany it.
Foreign tourists were the target.
The discussion became very lively.
The plan was decided to be implemented.

智美は夏には避暑地を訪れる。
今年も行きつけのコテージを訪れた。
部屋にはカーテンがそよいでいた。
Satomi visits a resort in summer.
(Sh) visited her favorite cottage again this year.
<table>
<thead>
<tr>
<th>232</th>
<th>Curtains were rustling in the room.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>‘Tomomi {was cooling/cooled} off by the window quietly.’</td>
</tr>
<tr>
<td></td>
<td>Wheatfields spread outside the window.</td>
</tr>
<tr>
<td></td>
<td>Then a cat appeared.</td>
</tr>
<tr>
<td></td>
<td>It was a cat with gray fur.</td>
</tr>
<tr>
<td></td>
<td>Tomomi held out her hand in an attempt to pet it.</td>
</tr>
<tr>
<td></td>
<td>When petting, the cat’s fur had static electricity.</td>
</tr>
<tr>
<td></td>
<td>Wheatfields spread outside the window.</td>
</tr>
<tr>
<td></td>
<td>Then a cat appeared.</td>
</tr>
<tr>
<td></td>
<td>It was a cat with gray fur.</td>
</tr>
<tr>
<td></td>
<td>Tomomi held out her hand in an attempt to pet it.</td>
</tr>
<tr>
<td></td>
<td>When petting, the cat’s fur had static electricity.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>15</th>
<th>Tatsuya lives in an apartment.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Last week, the bath broke.</td>
</tr>
<tr>
<td></td>
<td>Tatsuya went to the public bath after a long time.</td>
</tr>
<tr>
<td></td>
<td>Tatsuya felt a little dizzy.</td>
</tr>
<tr>
<td></td>
<td>Tatsuya felt like drinking coffee-flavored milk.</td>
</tr>
<tr>
<td></td>
<td>(He) began looking forward to a cup (of the drink) after bath.</td>
</tr>
<tr>
<td></td>
<td>(He) spilled it on his clothes drinking all at once after bath.</td>
</tr>
<tr>
<td></td>
<td>Tatsuya smiled wryly since he had just taken a bath.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>16</th>
<th>Atsushi is a college student.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Last night, Atsushi was up all night.</td>
</tr>
<tr>
<td></td>
<td>Since Atsushi was a morning person, it fatigued him.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>17</th>
<th>Kohei is a figure-skater.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>He has achieved remarkable growth.</td>
</tr>
<tr>
<td></td>
<td>Kohei was devoted to practicing today</td>
</tr>
<tr>
<td>康平は音楽にあわせて上手に踊っていた。/踊った。</td>
<td>as usual.</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>‘Kohei {was dancing/danced} well to the music.’</td>
<td>Kohei was superior in expression. Jumping ability was also important to win in the competition. It was a key to keep good balance between flexibility and muscle strength. Kohei decided to have a training coach. Kohei swore to himself that he would mount the podium this season.</td>
</tr>
<tr>
<td>康平は表現力に優れていた。競技会で勝つにはジャンプ力も大切だった。筋力と柔軟性をバランスよく保つのがコツだった。康平はトレーニングコーチを付けることにした。康平は今シーズンは表彰台にのぼると心に誓った。</td>
<td>Kohei was superior in expression. Jumping ability was also important to win in the competition. It was a key to keep good balance between flexibility and muscle strength. Kohei decided to have a training coach. Kohei swore to himself that he would mount the podium this season.</td>
</tr>
<tr>
<td>一輝はこのところ運動不足だった。彼はジョギングを日課にすることにした。朝の空気は新鮮だった。</td>
<td>Kazuki had not gotten enough exercise lately. He made it a daily routine to jog. The morning air was fresh.</td>
</tr>
<tr>
<td>一輝は川沿いを走っていた。/走った。</td>
<td>‘Kazuki {was running/ran} along the river.’</td>
</tr>
<tr>
<td>新緑が目に心地良かった。車の往来が少ないのも良かった。習慣にするとやめられないと思った。一輝は日増しに脚が強くなるのを感じた。K一輝は自分の変化に驚いた。</td>
<td>The fresh green color was pleasing to his eyes. It was good, too, that the traffic was less. (He) did not feel like quitting it once (he) made it a habit. Kazuki felt that his legs became stronger day by day. Kazuki was surprised at the change in himself.</td>
</tr>
<tr>
<td>久美子と稔は映画を見に出かけた。久美子が好きなコメディ映画を選んだ。二人は列の真ん中の席に着いた。</td>
<td>Minoru and Kumiko went to the movies. (They) chose Kumiko’s favorite, a comedy movie. They sat down in the middle of the row.</td>
</tr>
<tr>
<td>久美子は思う存分笑っていた。/笑った。</td>
<td>‘Kumiko {was laughing/laughed} to her heart’s content.’</td>
</tr>
<tr>
<td>ストレス発散にはもってこいだった。久美子は小腹が空いてきた。館内に売店があったのを思い出した。久美子は席をたった。久美子はうっかり観客の足を踏んでしまった。久美子は冷や汗をかいて平謝りした。</td>
<td>It was perfect for releasing stress. Kumiko became a little hungry. (She) remembered there was a shop on site. Kumiko got up from her seat. Kumiko stepped on an audience member’s foot inadvertently.</td>
</tr>
<tr>
<td>Page</td>
<td>Japanese Text</td>
</tr>
<tr>
<td>------</td>
<td>--------------</td>
</tr>
<tr>
<td>20</td>
<td>明美はホテルでアルバイトをしていた。それは結婚式場での配膳係だった。一日で一万円を稼ぐことができた。</td>
</tr>
<tr>
<td></td>
<td>その土曜日にも明美は朝から働くっていた。</td>
</tr>
<tr>
<td></td>
<td>午後には少し疲れが出てきた。コース料理の配膳は大変だった。明美はテーブルに汁物をこぼしてしまった。お客さんは気にしないようにと言ってくれた。披露宴ではお客さんが寛大なのが助かった。</td>
</tr>
<tr>
<td>21</td>
<td>舞は理恵とビーチにやって来た。それは海ガメが有名なビーチだった。海水は透きとおってさらさらしていた。</td>
</tr>
<tr>
<td></td>
<td>舞はのびのびと泳いでいた。</td>
</tr>
<tr>
<td></td>
<td>太陽の光が水面にきらきら輝いていた。遠くの浜辺にごそごそ動く影が見えた。理恵が、海ガメだと教えてくれた。舞と理恵は驚かさぬよう遠くから見守った。舞は次回は望遠カメラを持ってこようと思った。</td>
</tr>
<tr>
<td>22</td>
<td>順子は小学生だ。先週には珍しく大雪が降った。順子はさっそく外に飛び出した。</td>
</tr>
<tr>
<td></td>
<td>順子は庭で楽しく遊んでいた。</td>
</tr>
<tr>
<td></td>
<td>順子は雪ウサギを作った。固めの雪で、まとめるのに苦労した。しかしべタべタな雪よりはましだった。南天の実と葉で表情を作った。雪ウサギは日陰で何日も姿を留めていた。</td>
</tr>
<tr>
<td>23</td>
<td>The snow rabbit kept its figure for many days in the shade.</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>麻衣は春休みにタイに旅行した。</td>
<td></td>
</tr>
<tr>
<td>Mai traveled to Thailand during spring break.</td>
<td></td>
</tr>
<tr>
<td>タイは都市部が意外と近代的だった。</td>
<td></td>
</tr>
<tr>
<td>Thailand was surprisingly modern in urban areas.</td>
<td></td>
</tr>
<tr>
<td>そして夜のマーケットが有名だ。</td>
<td></td>
</tr>
<tr>
<td>And the night markets are famous.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>24</th>
<th>Yuko’s hobby is drawing picture-letters.</th>
</tr>
</thead>
<tbody>
<tr>
<td>裕子の趣味は絵手紙を描くことだ。</td>
<td></td>
</tr>
<tr>
<td>She favors the fruits for the theme.</td>
<td></td>
</tr>
<tr>
<td>彼女は果物を好んで題材にする。</td>
<td></td>
</tr>
<tr>
<td>Yuko has changed the mood that day.</td>
<td></td>
</tr>
<tr>
<td>裕子はその日は気分を変えた。</td>
<td></td>
</tr>
<tr>
<td>互いの絵を品評するのが楽しみだった。</td>
<td></td>
</tr>
<tr>
<td>(She) carefully expressed the texture of the petals.</td>
<td></td>
</tr>
<tr>
<td>花びらの質感を丁寧に表現した。</td>
<td></td>
</tr>
<tr>
<td>Yuko often sent a picture-letter to her grandmother.</td>
<td></td>
</tr>
<tr>
<td>裕子は絵手紙をいつも祖母に送っていた。</td>
<td></td>
</tr>
<tr>
<td>Her grandmother had adequate knowledge of painting.</td>
<td></td>
</tr>
<tr>
<td>祖母は絵に造詣があった。</td>
<td></td>
</tr>
<tr>
<td>It was fun to exchange evaluations on pictures with each other.</td>
<td></td>
</tr>
<tr>
<td>互いの絵を品評するのが楽しみだった。</td>
<td></td>
</tr>
<tr>
<td>祖母は今回も裕子の繊細な描写を褒めた。</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>23</th>
<th>Mai {was wandering/wandered} around the stalls.’</th>
</tr>
</thead>
<tbody>
<tr>
<td>麻衣は夜店をふらついていた。</td>
<td></td>
</tr>
<tr>
<td>(She) was able to enjoy delicious foods cheap.</td>
<td></td>
</tr>
<tr>
<td>/ぶらついた。</td>
<td></td>
</tr>
<tr>
<td>Mai ordered a noodle dish.</td>
<td></td>
</tr>
<tr>
<td>麻衣は夜店を（ぶらついていた。/ぶらついた。}</td>
<td></td>
</tr>
<tr>
<td>Herbs and peanuts gave zests to it.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>24</th>
<th>‘Yuko {was drawing/drew} a lily flower enthusiastically.’</th>
</tr>
</thead>
<tbody>
<tr>
<td>裕子は熱心にユリの花を{描いていた。/描いた。}</td>
<td></td>
</tr>
<tr>
<td>(She) carefully expressed the texture of the petals.</td>
<td></td>
</tr>
<tr>
<td>すりの花を{描いていた。/描いた。}</td>
<td></td>
</tr>
<tr>
<td>Yuko’s delicately depicted this time again.</td>
<td></td>
</tr>
<tr>
<td>裕子は熱心にユリの花を（描いていた。/描いた。}</td>
<td></td>
</tr>
<tr>
<td>Her grandmother praised Yuko’s delicate depiction this time again.</td>
<td></td>
</tr>
</tbody>
</table>
## APPENDIX H

### EXPERIMENTS 7 AND 9 (RESULTATIVE), CRITICAL SENTENCES

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Japanese</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>太郎は月曜日に教授と面談があった。太郎は資料を持参しようと考えた。太郎はまず資料をコピーしようとした。</td>
<td>Taro had a meeting with a professor on Monday. Taro thought to bring reference material. Taro first tried to copy the material.</td>
</tr>
<tr>
<td></td>
<td>すると運悪くコピー機が壊れていた。</td>
<td>‘Then unfortunately the copier {had been/was} completely broken.’</td>
</tr>
<tr>
<td></td>
<td>しかし太郎はすぐに機転を利かせた。太郎はその資料の PDF ファイルを探し始めた。太郎はそれをプリントアウトした。太郎の努力は教授に明確に伝わった。教授は太郎の論文に具体的な助言をくれた。</td>
<td>But Taro managed with tact immediately. Taro looked for a PDF file of the document. Taro printed it out. Thanks to it, (he) was able to prepare good materials. Efforts of Taro were conveyed to professor clearly. Professor gave him specific advice on Taro’s paper.</td>
</tr>
<tr>
<td>2</td>
<td>明美は短大に通っていた。明美は特に外国語に興味があった。彼女はフランス文学に傾倒した。明美は大学へ編入することに決まっていた。</td>
<td>Akemi had attended junior college. Akemi was especially interested in foreign languages. She committed to French literature. Akemi {had decided/decided} to transfer to the university.’</td>
</tr>
<tr>
<td></td>
<td>明美は日々ラジオでフランス語を聞いた。その経験は役立った。</td>
<td>Akemi listened to French on the radio daily. That experience helped.</td>
</tr>
<tr>
<td></td>
<td>明美は日々ラジオでフランス語を聞いた。その経験は役立った。</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>智美は図書館で働いている。その日、智美は本のラベルの張替えをdid.</td>
<td>Satomi works at the library. That day, Satomi re-covered labels on books.</td>
</tr>
<tr>
<td>行った。</td>
<td>First (she) peeled labels from old books.</td>
<td></td>
</tr>
<tr>
<td>まずは古い本からラベルを剥がした。</td>
<td>すると本にはラベルの跡がくっきり {残っていた。/残った。}</td>
<td></td>
</tr>
<tr>
<td>‘Then the mark of the label {had remained/remained} clearly on the book.’</td>
<td>本は少し日焼けをしていたようだ。</td>
<td></td>
</tr>
<tr>
<td>また、糊が完全には取り除けていなかった。</td>
<td>The book seemed a little sunburnt.</td>
<td></td>
</tr>
<tr>
<td>智美は表面をなるべく綺麗にした。</td>
<td></td>
<td></td>
</tr>
<tr>
<td>それから新しいラベルを貼った。</td>
<td>In addition, the glue was not completely removed.</td>
<td></td>
</tr>
<tr>
<td>智美は本が少し生き返った心地がした。</td>
<td>Satomi made the surface as clean as possible.</td>
<td></td>
</tr>
<tr>
<td>然后 (she) put a new label.</td>
<td>Satomi felt like the book was revived a little.</td>
<td></td>
</tr>
</tbody>
</table>

| 久美子は香川県出身だ。 | Kumiko comes from Kagawa Prefecture. |
| 久美子は一人っ子だった。 | Kumiko was an only child. |
| 彼女は五年前に大学を卒業した。 | She graduated from college five years ago. |
| 久美子はこの春に {結婚していた。/結婚した。} | ‘Kumiko {had been/was} married this spring.’ |
| 夫には姉がひとりいた。 | Her husband had one sister. |
| 久美子は義理の姉ができて喜んだ。 | Kumiko was delighted as she now had a sister-in-law. |
| 彼女達はすぐに仲良くなった。 | The women became friends immediately. |
| 元気な姉と二人で愉快に過ごした。 | Last week, Kumiko went to see her sister. |
| (She) had a good time together with her active sister. |

<p>| 裕子は夕飯の準備をしていた。 | Yuko was preparing dinner. |
| 夕飯は洋食に決めた。 | (She) decided on western dishes for dinner. |
| 日持ちのするシチューを煮込むことにした。 | (She) decided to simmer stew, which would stay good for days. |
| その夜は健太は勤務先の花見に出かけていた。/出かけた。 | ‘Kenta {had been/went} out for a cherry blossom viewing hosted by the employer that night.’ |
| 健太の留守中に、自宅に電話があった。 | In the absence of Kenta, there was a call at home. |
| それは直樹からの電話だった。 | It was a phone call from Naoki. |
| 健太に用事があるとのことだった。 | (He) wanted to contact Kenta. |
| 裕子には、健太がいつ帰宅するか分からなかった。 | Yuko did not know when Kenta would come back home. |
| 裕子は直樹にそう伝えた。 | Yuko told so to Naoki. |</p>
<table>
<thead>
<tr>
<th>セル</th>
<th>日本語</th>
<th>英語</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>直樹は明朝かけなおすと裕子に言った。</td>
<td>Naoki told Yuko that he would call again the next morning.</td>
</tr>
<tr>
<td></td>
<td>恵はハワイで暮らしている。海外生活はもう六年になる。毎年、冬はニューヨークに出かける。</td>
<td>Megumi lives in Hawaii. (She) has been living abroad six years. Every year, (she) goes out to New York in winter.</td>
</tr>
<tr>
<td></td>
<td>去年の冬は、気分を変えて日本に戻っていた。</td>
<td>‘(She) {had been/was} back to Japan last winter for a change.’</td>
</tr>
<tr>
<td></td>
<td>恵にはお正月の雰囲気が懐かしかった。その正月はことさら底冷えた。しかし身を切る寒さが新鮮だった。</td>
<td>(She) spent New Year’s holidays in Japan after a long time. Atmosphere of the New Year was nostalgic to Megumi. It was bitterly cold especially that New Year. However, the bitter cold was refreshing (to her). It was extra fun to get warm in the kotatsu.</td>
</tr>
<tr>
<td>7</td>
<td>美穂は小学校生だ。先週、赤い羽根募金が実施された。美穂もすすんで募金をした。</td>
<td>Miho is an elementary school student. Last week, Red Feather fundraising was carried out. Miho willingly made a donation, too.</td>
</tr>
<tr>
<td></td>
<td>それから美穂はカバンに赤い羽根をつけていた。</td>
<td>‘Then Miho {had worn/wore} a red feather on her bag.’</td>
</tr>
<tr>
<td></td>
<td>羽根はシールで留めるものだった。ある日、羽根が取れかかっていた。美穂は工夫して、カバンに刺繍で縫いとめた。それは白いカバンによく似合った。美穂は良い事をしたのと綺麗なのでて嬉しかった。美穂は次の機会にも募金しようと思った。</td>
<td>The feather was worn with a sticker. One day, the feather had loosened. Miho thought things out and stitched it on the bag as embroidery. It went well with the white bag. Miho was happy with doing good as well as the beauty (of the feather). Miho thought to make a donation at the next opportunity as well.</td>
</tr>
<tr>
<td>8</td>
<td>亮は研究者だった。彼は先週、学会に参加した。初日の夕方には懇談会に出席した。</td>
<td>Akira was a researcher. Last week, he took part in an academic conference. (He) attended the reception on the first evening.</td>
</tr>
<tr>
<td></td>
<td>会場には和洋さまざまな料理が並んでいた。</td>
<td>‘There {had been/were} various Japanese and Western cuisines lined up at the venue.’</td>
</tr>
</tbody>
</table>
亮は、好物のから揚げを皿にとった。ついでにサラダを少しとった。会場には、大勢の知り合いが居た。なかでも懐かしい恩師に再会できた。亮にはそれは望外の喜びだった。

Akira put fried chicken, which was his favorite, on a plate. Then (he) took a little salad. Finally, (he) took a little ham. There were many acquaintances in the venue. Above all (he) was able to reunite with (his) former teacher whom he missed. It was un hoped-for joy for Akira.

午前中、和彦はずっと忙しかった。和彦には朝食をとる時間もなかった。そこで昼は定食屋に出かけることにした。

Kazuhiko was busy all morning. Kazuhiko had no time to have breakfast. So (he) decided to go out for lunch at a diner.

和彦は仕事部屋を一時間ほど空けていた。その後、彼はすぐパソコンに向かった。

‘Kazuhiko {had left/left} his office for about an hour.’

+(その後、彼はすぐパソコンに向かった。) すると優子からメールが届いていた。書類作成のお願いだった。+和彦は快諾した。+和彦の頭は冴えていた。おかげで書類作成は30分で済んでしまった。

He then headed to the computer immediately. Then an e-mail had arrived from Yuko. (She) was asking for a favor concerning paperwork. Kazuhiko readily agreed. Kazuhiko’s mind was clear. Thanks to it, (he) was done with the paperwork in 30 minutes.

佳奈の小学校では始業式があった。全校生徒が体育館に集合した。壇上には新任の先生達と校長先生が居た。

There was an opening ceremony of the school term in the elementary school that Kana attended. All the students gathered in the gym. The principal and new teachers were on the stage.

校長先生はマイクの前に立っていた。/立った。

‘The headmaster {had stood/stood} in front of the microphone.’

校長先生は、先生達の紹介を始めた。新任の先生は男性三人、女性二人だった。みな良さそうな先生だった。始業式は一時間で終わった。生徒達は新学期が楽しみになった。

The principal started introducing teachers. There were three new male teachers and two new female teachers. All the teachers looked nice. Opening ceremony was over in an hour. New term became fun for students.

斉藤さんは出版会社に勤めた。彼は周囲からよく慕われていた。

Mr. Saito worked for a publishing company. He was loved by those around him.
| 12 | 哲也はフィギュアスケートの選手だ。 | Tetsuya is a figure-skater. |
| 12 | 哲也は練習に余念がない。 | Tetsuya devotes himself to practicing. |
| 12 | 彼は今日の練習内容を振り返った。 | He looked back at the contents of the practice that day. |
| 13 | 翔は先週、入院した。 | Last week, Sho was hospitalized. |
| 13 | 翔は思うより元気そうだった。 | Sho looked better than (she) thought. |
| 14 | 舞は先週、チューリップの鉢植えを買った。 | Last week, Mai bought a potted tulip. |

| 12 | 斉藤さんは定年で今年の三月にその会社を辞めていた。 | Mr. Saito had left the company due to retirement age this March. |
| 12 | 其の趣味はゴルフだった。 | Mr. Saito’s hobby was golfing. |
| 12 | 彼は旅も趣味だった。 | Traveling was also his hobby. |
| 13 | 翔は先週、入院した。 | Last week, Sho was hospitalized. |
| 13 | 原因は盲腸だった。 | The cause was the cecum. |
| 14 | 舞は先週、チューリップの鉢植えを買った。 | Last week, Mai bought a potted tulip. |

| 12 | 斉藤さんの趣味はゴルフだった。 | Mr. Saito’s hobby was golfing. |
| 12 | 旅行も趣味だった。 | Traveling was also his hobby. |
| 12 | 彼は息子夫婦とハワイに訪れた。 | He went to Hawaii with his son and his daughter-in-law |
| 12 | 彼はハワイでゴルフをするのを楽しみにしていた。 | Ladies enjoyed shopping, men enjoyed golfing, to their heart’s content. |
| 14 | 舞は先週、チューリップの鉢植えを買った。 | Last week, Mai bought a potted tulip. |

| 12 | 斉藤さんの定年で今年の三月にその会社を辞めていた。 | Mr. Saito had left the company due to retirement age this March. |
| 12 | 其の趣味はゴルフだった。 | Mr. Saito’s hobby was golfing. |
| 12 | 旅行も趣味だった。 | Traveling was also his hobby. |
| 12 | 彼は息子夫婦とハワイに訪れた。 | He went to Hawaii with his son and his daughter-in-law |
| 12 | 彼はハワイでゴルフをするのを楽しみにしていた。 | Ladies enjoyed shopping, men enjoyed golfing, to their heart’s content. |
| 14 | 舞は先週、チューリップの鉢植えを買った。 | Last week, Mai bought a potted tulip. |

<p>| 12 | 斉藤さんの趣味はゴルフだった。 | Mr. Saito’s hobby was golfing. |
| 12 | 旅行も趣味だった。 | Traveling was also his hobby. |
| 12 | 彼は息子夫婦とハワイに訪れた。 | He went to Hawaii with his son and his daughter-in-law |
| 12 | 彼はハワイでゴルフをするのを楽しみにしていた。 | Ladies enjoyed shopping, men enjoyed golfing, to their heart’s content. |
| 14 | 舞は先週、チューリップの鉢植えを買った。 | Last week, Mai bought a potted tulip. |</p>
<table>
<thead>
<tr>
<th>文字列</th>
<th>英語翻译</th>
</tr>
</thead>
<tbody>
<tr>
<td>舞はつぼみが開くのを日ごと待った。</td>
<td>This morning, tulips {had bloomed/bloomed} beautifully.</td>
</tr>
<tr>
<td>今朝、チューリップは見事に咲いていた。</td>
<td>The dark yellow petals were beautiful. Mai examined the flowering period of the tulip.</td>
</tr>
<tr>
<td>花弁は濃い黄色でとても綺麗だった。</td>
<td>Mai thought to enjoy it at best. Accordingly, Mai transferred the pot to the table.</td>
</tr>
<tr>
<td>舞はチューリップの開花期間を調べた。</td>
<td>Tomomi gathered with two friends after a long time. Entering the cafe, (they found) it was overall relaxed inside. First (they) got drink at the stand.</td>
</tr>
<tr>
<td>およそ一週間とのことだった。</td>
<td>They engaged in a lively conversation. When they noticed, it was about time for the last bus to arrive. They disbanded reluctantly.</td>
</tr>
<tr>
<td>舞はせいぜい楽しもうと思った。</td>
<td>It was more northwest than it would be in the summer. Kazuya suddenly felt uplifted. Kazuya stared at the rainbow for a while. He was almost late.</td>
</tr>
<tr>
<td></td>
<td>Tatsuya had a cousin.</td>
</tr>
<tr>
<td>Page</td>
<td>Text</td>
</tr>
<tr>
<td>-------</td>
<td>-------</td>
</tr>
</tbody>
</table>
| いとこの大学は先週から冬期休暇だった。
いとこは海外旅行に行く予定だった。 | The university that the cousin went to had winter vacation from last week. The cousin planned to go on a trip abroad. |
| 達也はいとこにスーツケースを貸していた。
‘Tatsuya {had lent/lent} a suitcase to his cousin.’ |  |
| いとこは五日間ほどで帰ってきた。
いとこは台湾に旅行していた。
物価が安くて食べ物が美味しいらしい。
いとこは達也にたくさん土産を買ってきてくれた。
カレーライス 達也は中でも月餅がたいへん気に入った。 | The cousin came back in about five days. The cousin had traveled to Taiwan. It is said that prices are low and the food is delicious there. Cousin bought a lot of souvenirs for Tatsuya. Tatsuya liked mooncakes very much among others. |
| 康平は姫のドライブが好きだった。
その朝も濃い霧の中を進んだ。
康平は慎重に運転した。 | Kohei liked driving on hills. (He) went through a dense fog that morning as well. Kohei drove carefully. |
| 霧は頂上付近ですっかり{消えていた。/消した。} ‘The fog {had been/was} disappeared near the summit.’ |  |
| そしてまぶしい朝日が広がった。
姫の向こうまで景色が見渡せた。
緑濃い山並みが綺麗だった。
康平はうっとうして車を停めた。
朝のしゅっとった空気は心地良かった。 | Then the bright morning sun spread out. (You) could see as far as the other side of the hill. Dark green mountains were beautiful. Kohei happily parked the car. The moist morning air was pleasant. |
| 一輝は電車に乗っていた。
すると友美から着信があった。
一輝はデッキに移動して電話を取った。 | Kazuki was riding on the train. Then there was an incoming call from Tomomi. Kazuki moved to the deck and got the phone. |
| 電波が不安定で友美的声が{聞こえづらくなっていた。/聞こえづらくなった。} ‘The reception was not good and Tomomi’s voice {had become/became} difficult to hear (on the phone).’ |  |
| 1時に会議があるらしかった。
あと二時間しかなかった。
最寄り駅まで一時間はかかる距離だった。
一輝は食事をする間に節約した。
昼食はキヨスクでおにぎりを買った。
会社に着くと、会議は7時からである | There was to be a meeting at 1 o’clock. There were only two hours left. It would take at least one hour to the nearest station. Kazuki saved time for meals. (He) bought rice balls for lunch at the kiosk. |
| 20 | 今年も梅雨の季節になった。  
最近は入梅が早い。  
雨は先週もずっと降り続いた。 | The rainy season has come this year.  
The rainy season comes quickly recently.  
It kept raining a lot last week. |
|  | しかし今週は雨はぱったり{やんでいた。/やんだ。}  
‗But this week rain {had stopped/stopped} completely.’ |  |
| 21 | その日は快晴だった。  
由佳は帰り道をゆっくり歩いた。  
そろそろ陽が落ちてきた。 | The weather was fine that day.  
Yuka walked slowly on the way home.  
Gradually the sun set. |
|  | 西の空は薄茜色に{染まっていった。/染まった。}  
‗The west sky {had dyed/dyed} in pale madder red.’ |  |
| 22 | 拓也は学生寮に住んでいた。  
その寮は大学に近かった。  
友人の雄太もその寮に住んでいた。 | Takuya lived in a student dormitory.  
The dorm was close to the university.  
Yuta’s friend also lived in the dorm. |
|  | 先週に雄太の部屋の鍵が{なくなっていった。/ななくった。}  
‗The key to Yuta’s room {had gone/went} missing last week.’ |  |
<table>
<thead>
<tr>
<th>雄太はしばらく食費を節約することにした。</th>
<th>Yuta decided to save money on groceries for a while.</th>
</tr>
</thead>
<tbody>
<tr>
<td>健一と絵美は映画を見に出かけた。 絵美が好きな恋愛映画を選んだ。 二人は列の真ん中の席に着いた。</td>
<td>Kenichi and Emi went to the movies. They chose Emi’s favorite, a romance movie. They sat down in the middle of the row.</td>
</tr>
<tr>
<td>若者がその通路側の席に座っていた。</td>
<td>‘A young man {had sat/sat} on the aisle seat.’</td>
</tr>
<tr>
<td>映画を観始めて一時間経った。 絵美は席をたった。 例の若者は居眠りをしていた。 絵美はうっかり若者の足を踏んでもった。 若者は大声をあげた。 絵美は冷や汗をかいて平謝りした。</td>
<td>It had been an hour since they started watching the movie. Emi got up from her seat. That young man was dozing. Emi stepped on his foot inadvertently. The young man shouted. Emi made profuse apologies</td>
</tr>
<tr>
<td>八重子は茶道を習っている。 彼女はその日、和服を着てお茶会に参加した。 八重子はお茶をたて、客に運ぼうとした。</td>
<td>Yaeko takes lessons in tea-ceremony. That day, she took part in a tea party in a kimono. Yaeko made some tea, and was carrying it to a guest.</td>
</tr>
<tr>
<td>八重子は着物の裾をうっかり踏んでいた。</td>
<td>‘Yaeko {had stepped/stepped} on the hem of kimono carelessly.’</td>
</tr>
<tr>
<td>次の瞬間、よろけてしまった。 しかしお茶はしっかりとたてられていた。 そして少量しか茶碗に入っていなかった。 おかげで八重子は何もこぼさずに済んだ。 八重子は作法にも気をつけなければならないと思った。</td>
<td>The next moment, (she) stumbled. But thick tea had been made. And there was only a small amount in the tea bowl. Thanks to it, Yaeko did not spill anything. Yaeko thought she should also mind her manners.</td>
</tr>
</tbody>
</table>
REFERENCES


