Spring 1-1-2015

Now you see it, now you don’t: The influence of lexical class, usage frequency, and constructional context on semantic simulation

Kevin Michael Gould

University of Colorado at Boulder, gouldk@colorado.edu

Follow this and additional works at: https://scholar.colorado.edu/ling_gradetds

Part of the Cognitive Psychology Commons, and the Linguistics Commons

Recommended Citation

Gould, Kevin Michael, "Now you see it, now you don’t: The influence of lexical class, usage frequency, and constructional context on semantic simulation" (2015). Linguistics Graduate Theses & Dissertations. 40.
https://scholar.colorado.edu/ling_gradetds/40

This Dissertation is brought to you for free and open access by Linguistics at CU Scholar. It has been accepted for inclusion in Linguistics Graduate Theses & Dissertations by an authorized administrator of CU Scholar. For more information, please contact cuscholaradmin@colorado.edu.
NOW YOU SEE IT, NOW YOU DON’T: THE INFLUENCE OF LEXICAL CLASS, USAGE FREQUENCY, AND CONSTRUCTIONAL CONTEXT ON SEMANTIC SIMULATION

by

KEVIN M. GOULD

B.A., University of Illinois, 2006

A thesis submitted to the
Faculty of the Graduate School of the
University of Colorado in partial fulfillment of the requirement for the degree of
Doctor of Philosophy
Department of Linguistics

2015
This thesis entitled:
Now you see it, now you don’t: The influence of lexical class, usage frequency, and
constructional context on semantic simulation
written by Kevin M. Gould
has been approved for the Department of Linguistics

_______________________
Laura A. Michaelis

_______________________
Bhuvana Narasimhan

Date____________

The final copy of this thesis has been examined by the signatories, and we
find that both the content and the form meet acceptable presentation standards
of scholarly work in the above mentioned discipline.

IRB protocol # 12-0243, 13-0412, 13-0653
ABSTRACT

Simulation semantics has successfully demonstrated that (1) linguistic knowledge is not relegated to one specific module, region, or network of the brain, (2) linguistic knowledge is built from an amalgam of embodied experiences, and (3) the simple act of reading a sentence may therefore evoke a distributed pattern of neurological activation that can interfere with or facilitate processing in (canonically) non-linguistic cognitive domains like, e.g., visual shape categorization or motor movement. These findings have had important ramifications for theories of linguistic representation, but also leave room for improvement in that they have not taken into account many factors central to functionalist linguistic description, including lexical class, usage frequency, and constructional context.

In this thesis I demonstrate within two different simulation modalities (motor simulation and a new paradigm called temporal simulation), across four experiments, that taking into account concepts of cognitive-functional linguistics like those listed above may help to explain the variable results often reported in simulation semantics studies. The concepts investigated include three forms of frequency (short-term frequency, long-term frequency, and verbal constructional ‘preference’), two kinds of lexical-class divisions (Vendler’s (1957) Aktionsart classes and Rappaport Hovav and Levin’s (2008) transfer verb classes), a construction-based opposition (ditransitive vs. oblique goal), and others. Ultimately, it is shown that these factors have the power to heavily distort and potentially even reverse simulation effects.

In order to capture the many effects of the above factors on mental simulation, in this thesis I develop and advocate for a model called sub-event monitoring, wherein the extent to which speakers inspect conceptual structure during simulation directly affects observed simulation results. A high degree of sub-event monitoring is associated with increased processing load and diminished effects of frequency on simulation, while a low degree of sub-event monitoring is associated with decreased processing load and enhanced effects of frequency on simulation. It will be shown that the vast majority of the data gathered in this dissertation can be subsumed within the sub-event monitoring framework, which can in turn be used to generate novel predictions about the effects of the above lexical and pragmatic factors on mental simulation.
ACKNOWLEDGEMENTS

A successful dissertation is not the product of the efforts of one individual, but is instead the product of years of scholarship and mentorship, collaboration and exploration, and, above all else, patience from an entire group of dedicated people. Of those who were part of this group, Laura Michaelis, my dissertation advisor and committee chair, certainly deserves to be recognized first and most emphatically. Her enthusiasm and support have helped me to find my identity as a linguist, and her preternatural ability to always recognize the bigger theoretical issues at stake has helped me to learn to appreciate the significance of my own work. I could not have possibly asked for a better mentor and friend over these past several years.

Thanks and recognition are also owed to the other members of my committee: Bhuvana Narasimhan, David Rood, Rebecca Scarborough, and Javier Rivas. I am indebted to Bhuvana, in particular, for her help in refining the methodological approaches and statistical interpretation employed in this dissertation. I’m equally indebted to Javier for his help in critiquing and improving upon the Spanish language items utilized in one of the experiments featured in this dissertation. Drs. Scarborough and Rood have, similarly, provided thorough and incisive feedback on this project, particularly during the dissertation proposal defense. I therefore owe them a debt of gratitude, as well.

Two additional members of the University of Colorado Linguistics Department need to be mentioned here. They are Barbara Fox and Steve Duman. Barbara served for a period as my co-advisor along with Laura, and managed to completely change my worldview of linguistics in a matter of weeks. She and I spent many hours in conversation pitting the linguistics of Noam Chomsky against those of Joan Bybee, the latter of which is featured prominently in this work. On a similar note, I should thank Steve Duman for not only being a good office mate off of
whom I could bounce even the most ridiculous methodological idea without any fear of judgment, but also for being a great friend and remarkable business partner.

Outside of the world of linguistics, there are five others whom I would be remiss not to thank here. I am sitting here writing this section thanks to the efforts of Drs. Tim Hain and Marcello Cherchi, who successfully treated a scary, enigmatic health issue I had a few years ago which would otherwise have certainly ended my aspirations of getting a doctorate in linguistics.

I need to thank my mom, Sally, for so many things that I couldn’t possibly list them here, and I need to thank my fiancée, Lisa, for her endless patience and support in dealing with an exhausted (and, very often, grouchy) man hunched over his laptop nearly every night for the better part of a year. Both of you mean the world to me.

Lastly, I should thank my dog, Pongo, who would always let me know that I had spent too much time writing by jumping up on my lap and licking my face.
CONTENTS

ABSTRACT.......................................................................................................................... iii

ACKNOWLEDGEMENTS........................................................................................................ iv

CONTENTS.......................................................................................................................................................... vi

LIST OF TABLES....................................................................................................................................................... xi

LIST OF FIGURES................................................................................................................................................... xii

CHAPTER 1. INTRODUCING LEXICO-PRAGMATIC SIMULATION SEMANTICS.......................................................................................................................... 1

1. Introduction...................................................................................................................................................... 1

1.1 Simulation Semantics................................................................................................................................. 2

1.2 On Abductive Inference in Linguistics and Cognitive Science................................................................. 4

1.3 Goals of the Dissertation............................................................................................................................ 7

1.4 Dissertation Plan........................................................................................................................................... 9

1.4.1 Experiment One (Chapter 3)................................................................................................................. 9

1.4.2 Experiment Two (Chapter 4).................................................................................................................. 10

1.4.3 Experiment Three (Chapter 5)............................................................................................................... 10

1.4.4 Experiment Four (Chapter 6)................................................................................................................ 11

1.5 Conclusions.................................................................................................................................................... 11

1.6 Looking Ahead to Chapter 2...................................................................................................................... 12

CHAPTER 2. KEY CONCEPTS........................................................................................................................ 14

2. Introduction.................................................................................................................................................. 14

2.1 Background on Simulation Semantics....................................................................................................... 14

2.1.1 Problems with Simulation Semantic Representation........................................................................... 17

2.2 Lexico-Pragmatic Factors and Sub-Event Monitoring............................................................................. 28
2.2.1 Usage Frequency ................................................................. 28
2.2.2 Lexical Class ........................................................................ 32
2.2.3 Temporal Density ................................................................. 35
2.2.4 Other Lexico-Pragmatic Factors ........................................ 41
2.2.5 Sub-Event Monitoring Explained ........................................ 41
2.3 Methodologies ........................................................................ 47
2.3.1 The Motor Facilitation and Interference Paradigm ............. 47
2.3.2 The Temporal Estimation Paradigm .................................... 52
2.4 Summary of Chapter 2 ............................................................ 60
2.5 Looking Ahead to Chapter 3 .................................................... 61

CHAPTER 3. SHORT-TERM FREQUENCY, GRAMMATICAL PATTERNS, AND MOTOR SIMULATION EFFECTS IN SPANISH ............................................... 62

3. Introduction .............................................................................. 62
3.1 Predictions ............................................................................. 64
3.2 Usage Frequency, Grammar, and Motor Simulation .................. 65
  3.2.1 The Motor Facilitation and Interference Paradigm .............. 66
  3.2.2 The Role of Grammar in Mental Simulations ................... 67
  3.2.3 Short-Term Frequency and Mental Simulation .................. 69
3.3 Methods ............................................................................... 70
  3.3.1 Participants ...................................................................... 74
  3.3.2 Materials ......................................................................... 74
  3.3.3 Design and Procedure .................................................... 75
3.4 Results .................................................................................. 77
3.5 General Discussion ............................................................... 85
3.6 Conclusions.................................................................................................................. 98
3.7 Looking Ahead to Chapter 4.......................................................................................... 100

CHAPTER 4. EXPLORING TEMPORAL SIMULATION EFFECTS..................................... 102

4. Introduction.................................................................................................................... 102
4.1 Time Perception and Temporal Density...................................................................... 103
    4.1.1 On Subjective Time, Objective Time, Temporal Dilation, and Temporal
          Compression.............................................................................................................. 105
4.2 Predictions.................................................................................................................... 107
4.3 A Note on Temporal Simulation Effects and Morphosyntactic Complexity.......... 110
4.4 Methods....................................................................................................................... 112
    4.4.1 Participants........................................................................................................... 112
    4.4.2 Materials.............................................................................................................. 113
    4.4.3 Design and Procedure......................................................................................... 115
4.5 Results......................................................................................................................... 117
4.6 General Discussion...................................................................................................... 122
4.7 Conclusions................................................................................................................ 134
4.8 Looking Ahead to Chapter 5....................................................................................... 136

CHAPTER 5. AKTIONSART CLASS, USAGE FREQUENCY, AND TEMPORAL
SIMULATION EFFECTS........................................................................................... 138

5. Introduction.................................................................................................................... 138
5.1 Predictions.................................................................................................................... 141
5.2 Frequency, Time Perception, and Aktionsart Class.................................................... 142
    5.2.1 Usage Frequency.................................................................................................. 143
    5.2.2 Information Complexity, Expectedness, and Time Perception.......................... 144
6.3.2 Materials........................................................................................................... 190

6.3.3 Design and Procedure.......................................................................................193

6.4 Results...................................................................................................................194

6.5 General Discussion..............................................................................................200

6.6 Conclusions...........................................................................................................211

6.7 Looking Ahead to Chapter 7................................................................................214

CHAPTER 7. GENERAL DISCUSSION AND CONCLUSION............................................ 216

7. Simulation Semantics through the Lens of Cognitive-Functionalist Linguistics (and
Vice Versa) ................................................................................................................. 216

7.1 Revisiting the Goals of the Dissertation............................................................. 217

7.2 Mental Simulation and Sub-Event Monitoring.................................................... 221

7.3 On Differential Frequency Effects........................................................................ 230

7.4 General Conclusions........................................................................................... 233

7.5 Possibilities for Future Study................................................................................ 237

7.6 A Final Thought.................................................................................................... 240

REFERENCES.............................................................................................................. 241

APPENDIX 1: Critical Stimuli Used in Experiment One............................................. 252

APPENDIX 2: Critical Stimuli Used in Experiment Two............................................. 255

APPENDIX 3: Critical Stimuli Used in Experiment Three........................................... 257

APPENDIX 4: Critical Stimuli Used in Experiment Four............................................. 259
LIST OF TABLES

TABLE 2.1: Lexico-pragmatic features implicated in sub-event monitoring to be featured in this dissertation................................................................................................................................. 47

TABLE 3.1: Spanish verbs and their paired objects used in experiment one................................................................. 70

TABLE 4.1: Fast-speed and slow-speed verbs and their frequencies................................................................. 115

TABLE 4.2: Fast-speed/slow-speed verb phrases and their grammatical aspects................................................................. 129

TABLE 5.1: Average COCA corpus frequencies for each Aktionsart class category used in experiment three......................................................................................................................................... 151

TABLE 5.2: Average COCA corpus frequencies for different ‘expectedness’ conditions used in experiment three......................................................................................................................................... 152

TABLE 6.1: Dative-alternating verb classes and constructional their ‘preferences’ utilized in experiment four......................................................................................................................................... 191

TABLE 7.1: Interplay between markedness effects in sub-event monitoring observed in experiment four......................................................................................................................................... 225

TABLE 7.2: Summary of experimental results featured in this dissertation, along with how theses results factor into the theory of sub-event monitoring......................................................................................................................................... 229
LIST OF FIGURES

FIGURE 2.1: Variation in reported response times (and effects) in simulation semantics studies.......................................................................................................................... 18

FIGURE 2.2: Simulation effects construable as either facilitation or interference (reprinted from Glenberg and Kaschak (2002)).............................................................................................................. 21

FIGURE 2.3: Simulation effects construable as either facilitation or interference (reprinted from Tseng and Bergen (2005)).............................................................................................................. 22

FIGURE 2.4: An unambiguous simulation effect within the visual modality (reprinted from Bergen (2005)).......................................................................................................................... 23

FIGURE 2.5: An example of inconsistent patterning in motor simulation results (reprinted from Bergen and Wheeler (2010)).......................................................................................... 24

FIGURE 2.6: An illustration of transitions in the representational spaces of three different predications.............................................................................................................................. 37

FIGURE 2.7: The hypothetical effect of a manner adverb on temporal density.......................................................................................................................... 38

FIGURE 2.8: The hypothetical effect of temporal scope on temporal density.......................................................................................................................... 39

FIGURE 2.9: A hypothetical rendering of the relationship between frequency and temporal density in determining simulation effects.......................................................................................................................... 45

FIGURE 2.10: A diagram of the different tiers of sub-event monitoring.......................................................................................................................... 46

FIGURE 2.11: A photograph of a typical motor facilitation/interference setup.......................................................................................................................... 49

FIGURE 2.12: The effect of response cue timing on motor simulation effects (reprinted from Borreggine and Kaschak (2006)).............................................................................................................. 51

FIGURE 2.13: A simple illustration of the progression of a single trial in the temporal estimation paradigm.......................................................................................................................... 55

FIGURE 2.14: A breakdown of how different temporal estimation methodologies require different interpretations.................................................................................................................. 59

FIGURE 3.1: A significant correlation between verb mention number and participant response time.......................................................................................................................... 78

FIGURE 3.2: A significant two-way interaction between verb type and sentence congruity.......................................................................................................................... 79
FIGURE 3.3: A reframing of the interaction featured in Figure 3.2 as an interaction between button press location and verb congruity................................................................. 80

FIGURE 3.4: A significant two-way interaction between grammatical pattern and verb mention number................................................................. 81

FIGURE 3.5: Relationship between grammatical pattern and grammatical pattern mention number.................................................................................. 82

FIGURE 3.6: A significant three-way interaction between grammatical pattern, verb congruency, and verb direction................................................................. 83

FIGURE 3.7: Significant differences between the morphological future and the simple present, present progressive, and preterite................................................................. 84

FIGURE 3.8: A significant three-way interaction between verb congruency, verb direction, and number of mentions......................................................................... 85

FIGURE 3.9: A significant three-way interaction between verb congruency, verb direction, and number of mentions (reprint of Figure 3.8)................................................................. 88

FIGURE 3.10: A significant three-way interaction between grammatical pattern, verb congruency, and verb direction (reprint of Figure 3.6)................................................................. 91

FIGURE 3.11: The effect of number of verb mentions of average user response time (sorted by individual verbs).................................................................................... 95

FIGURE 3.12: Relationship between grammatical pattern and grammatical pattern mention number (reprint of Figure 3.5)........................................................................ 96

FIGURE 4.1: Theoretically allowable relationships between average user response times for more temporally dense and less temporally dense items................................................................ 109

FIGURE 4.2: Adjusted average deviation from target response time sorted by stimulus sentence grammatical features........................................................................ 118

FIGURE 4.3: A significant effect for participant ADHD status in experiment two.................. 119

FIGURE 4.4: A significant effect for main verb usage frequency in fast-speed vs. slow-speed trials........................................................................................................ 120

FIGURE 4.5: A significant two-way interaction between main verb usage frequency and verb speed........................................................................................................ 121

FIGURE 4.6: Non-significant critical contrasts from experiment two (adjusted deviation)........................................................................................................ 123
FIGURE 4.7: Two pairs of critical contrasts from experiment two observed to significantly differ (adjusted deviation)............................................................................................................. 126

FIGURE 4.8: A significant effect for main verb usage frequency in fast-speed vs. slow-speed trials (reprint of Figure 4.4)........................................................................................................................................... 130

FIGURE 4.9: A significant two-way interaction between main verb usage frequency and verb speed (reprint of Figure 4.5)............................................................................................................................................... 132

FIGURE 5.1: Significant critical contrasts from experiment two...................................................... 140

FIGURE 5.2: A significant main effect for predication expectedness.............................................. 155

FIGURE 5.3: A significant main effect for subject frequency............................................................... 156

FIGURE 5.4: A significant main effect for Aktionsart class............................................................... 157

FIGURE 5.5: A significant main effect for verb frequency............................................................... 158

FIGURE 5.6: A significant two-way interaction between Aktionsart class and verb frequency....................................................................................................................................................... 159

FIGURE 5.7: Non-significant relationship between number of syllables and average deviation from target time....................................................................................................................................................... 161

FIGURE 5.8: Average deviation per verb by verb COCA corpus frequency..................................... 162

FIGURE 5.9: A significant main effect for predication expectedness (reprint of Figure 5.2)................................................................................................................................................................. 163

FIGURE 5.10: A significant main effect for Aktionsart class (reprint of Figure 5.4)....................... 165

FIGURE 5.11: Average response times for each Aktionsart class adjusted for syllable count and main verb frequency effects....................................................................................................................................................... 166

FIGURE 5.12: A significant two-way interaction between Aktionsart class and verb frequency (reprint of Figure 5.6)....................................................................................................................................................... 167

FIGURE 6.1: A significant effect for number of main verb mentions............................................. 195

FIGURE 6.2: A direction relationship between participant age and response time that approached significance................................................................................................................................................................. 196

FIGURE 6.3: Average participant response times for congruent and incongruent conditions................................................................................................................................................................. 197
FIGURE 6.4: A significant two-way interaction between construction type and sentence congruity.................................................................................................................. 198

FIGURE 6.5: A significant three-way interaction between construction type, sentence congruity, and verbal constructional ‘preference’.................................................................................................................. 199

FIGURE 6.6: A significant three-way interaction between construction type, sentence congruity, and verbal lexical class.................................................................................................................. 200

FIGURE 6.7: A significant two-way interaction between construction type and sentence congruity (reprint of Figure 6.4).................................................................................................................. 202

FIGURE 6.8: A significant three-way interaction between construction type, sentence congruity, and verbal constructional ‘preference’ (reprint of Figure 6.5).................................................................................................................. 203

FIGURE 6.9: A significant three-way interaction between construction type, sentence congruity, and verbal lexical class (reprint of Figure 6.6).................................................................................................................. 207

FIGURE 7.1: A significant three-way interaction between construction type, sentence congruity, and verbal constructional ‘preference’ featured in experiment four.................................................................................................................. 223

FIGURE 7.2: A significant three-way interaction between construction type, sentence congruity, and verbal lexical class featured in experiment four.................................................................................................................. 224

FIGURE 7.3: Two pairs of critical contrasts from experiment two observed to significantly differ.................................................................................................................. 226

FIGURE 7.4: A significant main effect for predication expectedness featured in experiment three.................................................................................................................. 226

FIGURE 7.5: A significant two-way interaction between Aktionsart class and verb frequency featured in experiment three.................................................................................................................. 228

FIGURE 7.6: A depiction of the various differential frequency effects repeatedly encountered throughout this dissertation.................................................................................................................. 231
Chapter 1: Introducing Lexico-Pragmatic Simulation Semantics

“Repetition is a form of change.”
– Brian Eno & Peter Schmidt, the Oblique Strategies

“The simulacrum is never that which conceals the truth – it is the truth which conceals that there is none. The simulacrum is true.”
– Jean Baudrillard, Simulacra and Simulation

1. Introduction

This dissertation is about the perceptually grounded cognitive representations, or ‘mental simulations,’ that language may evoke. Specifically, it investigates the relationship between mental simulations and certain foundational ideas of cognitive and functional linguistics which have heretofore not been sufficiently investigated or felicitously integrated within simulation semantics. These foundational ideas include, for example, usage frequency (in its various manifestations) and lexical class.

In investigating the relationship between cognitive-functional linguistics and mental simulation, many new discoveries about the nature and behavior of mental simulation are brought to light. Specifically, it will be shown here that mental simulation is far from an immutable cognitive phenomenon. Rather, the character of mental simulation changes due to a confluence of lexical and pragmatic factors, many of which can be subsumed under the theory of sub-event monitoring. The theory of sub-event monitoring states that when we simulate the content of a sentence, certain lexical and pragmatic factors cause us to ‘inspect’ simulation content more closely or less closely, affecting how simulation plays out. In subsequent chapters it will be shown that these lexico-pragmatic factors may interact with each other in complex ways, affecting the
process of sub-event monitoring to the extent that the relative configuration of results gathered in a simulation study can be amplified, nullified, distorted, or possibly even reversed.

In short, it will be demonstrated here that lexico-pragmatic factors strongly affect language-induced mental simulation, such that future studies should without question take these factors into consideration. Several other secondary findings will be provided throughout, including, most notably, the discovery of a new modality of language-induced mental simulation called temporal simulation.

The remaining sections of this chapter briefly detail the overarching concepts, goals, and conclusions of this dissertation. In the following section (Section 1.1), I provide some brief background on simulation semantics. After that, Section 1.2 describes the role of abductive inference in the theory-building presented in this dissertation, with special attention to its pitfalls and those measures taken to avoid such pitfalls. Next, the general goals of the dissertation are provided in Section 1.3, immediately followed by a brief summary of the dissertation’s four-experiment plan in Section 1.4. After that, Section 1.5 provides a brief summary of the conclusions that will be reached in this work. Finally, Section 1.6 previews the content to follow in Chapter 2.

1.1 Simulation Semantics

The field of simulation semantics came into being in the late 1990s with the publication of Barsalou’s (1999) ‘Perceptual Symbol Systems.’ In it, Barsalou argues that the predominant ways of representing conceptual knowledge in 20th century psychology and linguistics are fundamentally flawed. Specifically, he criticizes the so-called ‘amodal symbol systems’ of both fields of study. Broadly, amodal symbols systems are “representational schemes that [are] inherently nonperceptual” (ibid, p. 578). In these schemes, perceptual input is transduced to
amodal, abstract features. For example, using this means of description, the lexical entries for *sunset*, *tangerine*, and *traffic cone* would all share an abstract [+orange] feature which would link these items together in the lexicon, along with other predominantly orange things. Crucially, however, it needs to be pointed out that there is, according to this model, nothing inherently orange about an amodal symbol like [±orange]. Rather, [±orange] is best conceptualized as a variable that could capture any discrete perceptual feature (assuming here, of course, that discrete perceptual features exist), like [±imaginary], [±tall], or even, theoretically, [±flabbergasted].

What Barsalou (1999) suggests, ultimately, is that this step of transduction between perceptual input and amodal representation is superfluous. Instead, he argues, perceptual input (or subsets thereof) can be directly stored in long-term memory for use in various forms of mental representation. Stated differently, when we, for example, activate a word or phrase, we do not access an abstract lexical entry composed of amodal symbols, but instead access an amalgam of all of the sensory experiences associated with previous perceptually-grounded encounters with said word or phrase.

Several predictions follow from this claim. Firstly, if representation is not amodal, then the perception or use of certain linguistic concepts would not simply activate areas of the brain primarily dedicated to linguistic processing (e.g., Broca’s area, Wernicke’s area, etc.), but would additionally activate neural regions considered to be primarily dedicated to, say, motor control or visual cognition. Secondly, entire linguistic predications may therefore elicit complex and distributed patterns of neural activation based on a number of different factors (including, most pertinently for our purposes here, lexico-pragmatic factors like frequency and lexical class).

Since Barsalou’s seminal paper, there have been published several studies which indeed demonstrate that language is associated with patterns of neurological co-activation in diverse regions of the brain (Hauk, Johnsrude, & Pulvermüller, 2004; Pulvermüller, Hauk, Nikulin, &
These findings have changed how linguistic knowledge is conceptualized. Whereas previously predominant views have proposed an abstract, feature-based lexicon (see, e.g., Katz & Fodor, 1963 (Section 7); Fodor & Pylyshyn, 1988 (Section 2.1.3)), or a discrete ‘black box’ of linguistic competence or parameters to be set through basic exposure to language (see, e.g., Chomsky, 1965, 1981), a simulation-based model requires only linguistic and perceptual input in order to produce rich semantic representations.

For all of the growth and maturation of this representational model over recent years, however, there are still several areas where simulation semantics runs into problems. Chiefly, while simulation semantics experiments reliably generate significant results, these results may often conflict across experiments, yielding a framework that has yielded several inconsistent findings (the likes of which will be fully explicated in Chapter 2). What will be demonstrated in this dissertation is that approaching mental simulation with lexical and pragmatic factors in mind can potentially help to explain some of the inconsistency mentioned above.

1.2 On Abductive Inference in Linguistics and Cognitive Science

One way in which this dissertation represents a break from much of the previous work within cognitive linguistics is its use of abductive reasoning to unify lexical semantic and usage factors with simulation semantics. Abductive reasoning, simply put, is the formulation of a hypothesis or theory from observation (e.g., collecting data in a laboratory experiment). This is similar to the ‘bottom-up’ approach of induction, but as Deutscher (2002) points out, involves a further step of attributing a causal factor to an observed pattern. As an example, a linguist studying several different languages from a particular mountainous region in South America may
notice that all of these languages tend to lack nasal consonants. From this, he or she might reason that all of the world’s languages spoken in mountainous regions lack nasal consonants because the lack of oxygen forces speakers to breathe orally rather than nasally.

This type of reasoning is, admittedly, flawed in two separate areas. The first flaw is often referred to as Hume’s problem of induction, summarized by the following quotation (from Krueger, 2001, p. 16):

The supposition that the future resembles the past is not founded on arguments of any kind, but is derived entirely from habit, by which we are determined to expect for the future the same train of object to which we’ve been accustomed. (Hume, 1739, p. 184)

Returning to our example above, our hypothetical linguist, looking at several different languages spoken in mountainous regions and lacking nasal consonants, has potentially incorrectly inferred that all languages spoken in mountainous regions lack nasal consonants. Upon studying a new set of languages from another mountainous region, he or she might find a language with a rich inventory of nasal consonants and would therefore have to revise his or her working model in order to reflect that.

The second flaw above comes from the fact that our hypothetical linguist has attributed a causal explanation to an observed correlation which may, in fact, be wrong. In our hypothetical scenario, several languages may share a common feature (i.e., a lack of nasal consonants) for reasons other than effects of high-altitude respiratory effort. These languages could, for example, be genetically related. Alternatively, there may simply be no underlying cause for the observed cross-linguistic similarity. In this case, the perceived similarity may simply be due to chance.

In spite of the above pitfalls, there are several academic disciplines which regularly make use of abductive reasoning. Within cognitive science, for instance, abduction is used particularly frequently in fMRI studies. Poldrack (2006) summarizes the predominant line of reasoning below:
(1) In the present study, when task comparison A was presented, brain area Z was active.
(2) In other studies, when cognitive process X was putatively engaged, then brain area Z was active.
(3) Thus, the activity of area Z in the present study demonstrates engagement of cognitive process X by task comparison A. (p. 59)

To illustrate this line of reasoning, one might show pictures of cats and dogs to participants in an fMRI scanner and find that, when dog pictures are displayed, the basal ganglia are more active than when cat pictures are displayed. Citing previous research that has shown that the basal ganglia are active when people experience feelings of romantic love, the experimenters might therefore conclude that, relative to cats, dogs cause increased feelings of romantic love in humans.

Needless to say, such conclusions may be readily overturned. After running a few more participants, the researchers in our hypothetical fMRI study might find that a different brain area is actually more active than the basal ganglia when people see pictures of dogs (exemplifying Hume’s problem of induction). In addition, it may be that any number of the other functions of the basal ganglia (e.g., timekeeping, controlling eye movements, etc.) could be responsible for the observed result.

Like the hypothetical studies mentioned above, this dissertation employs some degree of abductive inference. However, I make a concerted effort to mitigate the classic pitfalls discussed previously. In particular, while experiments one and two (Chapters 3 and 4, respectively) represent experiments that rely heavily on abduction, the generalizations and theories which form this abduction are tested in experiments three and four (Chapters 5 and 6). In this regard, deductive reasoning is employed here in concert with abductive reasoning in order to increase the validity of the results reported.
Ultimately, it needs to be acknowledged here that the data and explanatory model presented in this dissertation are not the final word on the relationship between lexico-pragmatic factors and language-induced mental simulation. Rather, as more data become available in the future, it should be expected that this model will be reshaped and refined in order to better reflect the true behavior of linguistic representation.

1.3 Goals of the Dissertation

The first major aim of this dissertation is to show that certain lexico-pragmatic features affect language-induced mental simulation. The lexico-pragmatic feature of most import here is usage frequency, in its various manifestations (e.g., type vs. token, short-term vs. long-term, etc.). As Chapter 2 will make clear, usage frequency has far-reaching effects on how language is both used and interpreted. Therefore, in all four experiments (outlined immediately below in Section 1.4), some form of usage frequency serves as a predictor variable. Secondarily, in experiments three and four (Chapters 5 and 6), lexical class also serves as a predictor variable. The lexical classes featured are independently motivated, i.e., posited based upon the patterning of various verbs in spoken and written language. Other lexico-pragmatic factors are also examined (particularly in experiments one and two). These factors include tense, aspect, and modality marking (experiments two and three) and temporal density, a construct which captures the number of transitions in a predication’s event representation (experiments two and three).

After establishing the importance of lexico-pragmatic factors in shaping mental simulation, two secondary goals immediately present themselves. Specifically, after establishing a link between lexico-pragmatic features and mental simulation effects in the motor simulation paradigm, I will introduce a new simulation paradigm (temporal simulation), and examine these same factors in order to verify that our initial findings are not specific to one particular
experimental modality. Additionally, I conduct follow-up experiments in both the motor and temporal paradigms in order to provide some measure of evidence that the paradigm-internal results reported were, in fact, valid to begin with.

The final major goal of this dissertation is the development of a model that can (a) explain how certain lexico-pragmatic factors will affect mental simulation, and (b) predict what results one would expect to find under various experimental conditions and circumstances utilizing various lexico-pragmatic factors. Here, it will be shown that the various factors mentioned above can be subsumed within a new framework called sub-event monitoring, which serves to bridge lexico-pragmatic factors and mental simulation effects. The basis, character, and utility of sub-event monitoring will be highlighted in Chapter 2 and regularly revisited over the course of this dissertation (particularly in Chapter 7).

In light of the above, the goals of this dissertation can be summarized as follows:

(1) Establish a link between certain lexico-pragmatic factors (in particular, usage frequency and lexical class) and simulation-based representation.

(2) Show that this link exists in multiple simulation paradigms (motor simulation and temporal simulation) and in multiple languages (English and Spanish).

(3) Conduct follow-up experiments in both paradigms in order to verify the findings in 1 and 2, above.
Finally, develop a theory (sub-event monitoring) that successfully captures the patterning of all of the data gathered and that, ideally, will be able to predict the distribution of data in future experiments.

These goals will be revisited in the concluding chapter of this dissertation (Chapter 7), and discussed there in detail.

1.4 Dissertation Plan

This dissertation is composed of four experiments which collectively suggest that lexico-pragmatic factors have far-reaching effects in mental simulation. As stated above, all four experiments look at usage frequency to some extent. These experiments are performed using two different methodologies (the motor facilitation and interference paradigm and the temporal estimation paradigm), and each methodology is seen twice. This larger experimental configuration therefore allows for a certain degree of within-methodology replication as well as a certain degree of between-methodology replication. Each of the four experiments is summarized below in Sections 1.4.1-1.4.4.

1.4.1 Experiment One (Chapter 3)

Experiment one looks at the effects of short-term frequency in a motor simulation task using Spanish language clauses. While all of the critical trials in this depict manual motor action in some way (through verbs like lanzar (to throw) and beber (to drink)), these clauses vary in terms of their tense/aspect/modality combinations. Critically, the experiment uses a repeated-measures design wherein certain verb-object combinations are repeatedly encountered over the course of the experiment. These repeated exposures are viewed as a proxy for short-term frequency, and it
is hypothesized that as a verb-object combination’s short-term frequency changes, the extent to which it is inspected in sub-event monitoring will change as well, altering recorded simulation effects. Secondarily, it is also hypothesized that the various tense/aspect/modality combinations featured in experiment one will also be subject to different degrees of inspection in sub-event monitoring, and will each, therefore, yield differing simulation effects.

1.4.2 Experiment Two (Chapter 4)

Experiment two introduces a new experimental methodology to the field of simulation semantics—the temporal estimation paradigm. In this experiment, pairs of clause types of differing temporal density (a concept which will be discussed more thoroughly in Chapter 2) are compared against one another in order to see which clause pairs may yield significantly different temporal simulation effects. In this experiment, several critical clauses analyzed are also tagged for their main verb’s usage frequency in the COCA corpus (Davies, 2008-). In this case, it is hypothesized that there will be a complex interplay between different lexico-pragmatic features (i.e., usage frequency and temporal density) that will produce differential long-term frequency effects between conditions. This complex interaction effect will be introduced in Chapter 2 and explored in Chapter 7.

1.4.3 Experiment Three (Chapter 5)

Experiment three has several goals. Its main purpose is to test various Aktionsart classes within the temporal estimation paradigm in order to see if clauses differing in the lexical aspect of their main verb (rather than, say, clauses differing in their grammatical aspect) yield different temporal simulation effects. In this regard, experiment three represents both a methodological replication of experiment two and an exploration of lexical-class effects in mental simulation.
Additionally, all of the critical trials in experiment three track main verb frequency using frequency benchmarks from the COCA corpus.

A secondary set of critical trials is included in experiment three, as well. In this case, critical trials are designed to reflect expected and unexpected events (e.g., *a chef cooking pasta* versus *a chef cooking a sock*), and it is hypothesized that unexpected events will be inspected more closely than expected events in sub-event monitoring, therefore leading to differences in temporal simulation effects.

1.4.4 Experiment Four (Chapter 6)

Experiment four serves to conclude the course of research featured in this dissertation by looking at lexical class effects and frequency effects within the motor simulation paradigm. Specifically, Rappaport-Hovav and Levin’s (2008) dative-alternating verb classes are used in clauses denoting acts of transfer in order to verify (a) that lexical-class-based differences manifest themselves in differing motor simulation effects, and (b) that a given verb’s constructional ‘preference’ may manifest itself in differing motor simulation effects, as well. In this regard, it serves to partially replicate the motor simulation effects seen in experiment one, as well as replicate the lexical class effects seen in experiment three, albeit within a different methodology.

1.5 Conclusions

I will offer the following conclusions, summarized here and more extensively discussed in Chapter 7:

(1) Lexico-pragmatic factors (including usage frequency, lexical class, and others) strongly affect language-induced mental simulation. These factors can amplify, nullify, distort, or
potentially even reverse the configuration of average participant response times in a simulation-based study, and may even explain why certain experimental procedures within simulation semantics have seen inconsistent results. In other words, mental simulation effects cannot be derived from, say, lexical semantics alone. Rather, a combination of semantic and lexico-pragmatic factors needs to be taken into account in order to capture the complexity of the cognitive processes leading to simulation effects.

(2) The theory of sub-event monitoring can help us both to understand and predict the often complex and variable results seen in simulation semantics studies. In this theory, different types of lexico-pragmatic factors interact in order to yield results that would not have reached significance in studies that failed to control for these factors. In short, sub-event monitoring serves as the first instantiation of a new theory which links cognitive-functionalist linguistics with simulation semantics.

(3) Lastly, temporal simulation appears to be a valid, measurable form of mental simulation, and its associated methodology, the temporal estimation paradigm, appears to be valid, as well. This has exciting implications for simulation semantics: it not only opens up a new stream of research, but also suggests that other, less traditional forms of embodied cognition and representation (e.g., sense of equilibrium, emotional cognition, etc.) may be subject to experimental manipulation in future studies.

1.6 Looking Ahead to Chapter 2

The following chapter (Chapter 2) offers a broad conceptual overview of the considerations that inform this dissertation, ranging from the historical, to the theoretical, to the
methodological. This chapter includes not only a summary of the history of simulation semantics and its major findings, but also a summary of the past shortcomings of simulation semantics, several of which this dissertation hopes to address. This chapter also provides background on cognitive-functional linguistics, describing the variegated effects of frequency, lexical class, and other factors on linguistic structure, form, and representation.

Chapter 2 also describes the two major methodologies to be used in this dissertation, the motor facilitation and interference paradigm and the temporal estimation paradigm. As we will see, the motor facilitation and interference paradigm has a comparatively long but somewhat checkered background, while the temporal estimation paradigm represents a new experimental methodology within simulation semantics.

Finally, Chapter 2 thoroughly outlines two theoretical constructs that are central to the themes and conclusions of this dissertation: temporal density, a semantically-motivated construction which aims to capture the variability of temporal meaning in language (and which will come into play primarily in experiments two and three (Chapters 4 and 5)), and sub-event monitoring, the central theoretical model of this dissertation, which serves to bridge the gap between the various lexico-pragmatic factors mentioned above and the highly variable results of various simulation semantics studies.
Chapter 2: 
Key Concepts

2. Introduction

The major argument advanced in this thesis is that much of the variability seen in simulation semantic effects is attributable to perceptual properties of conceptual structures, in particular, the manner in which speakers inspect conceptual structure during sub-event monitoring in mental simulation. In order to make this case, I combine three models that connect linguistic cognition and linguistic form to human experience. The first, simulation semantics, is based on the use of embodied representations during sentence interpretation (Section 2.1). The second, the usage-based model, is based on the routinization of linguistic structures through repeated encounters (Section 2.2.1). The third, lexical semantics, is based on the verbalization of stereotypical patterns of linguistic and extralinguistic experience, as in the Aktionsart classes described by Vendler (1957) and Rappaport-Hovav and Levin (2008) (Section 2.2.2). Following these summaries, I outline the theory of sub-event monitoring, the central theoretical construct of this thesis, and describe the major predictions that it makes (Section 2.2.5). Lastly, I discuss the experimental methodologies employed in testing these predictions (Section 2.3). This conceptual overview should provide the reader with sufficient background information to understand both the major issues at stake in this dissertation, as well as the empirically driven means through which these major issues will be investigated.

2.1 Background on Simulation Semantics

A relatively new field of study, simulation semantics advocates for an inherently perceptual and image-based view of linguistic cognition (see, e.g., Barsalou, 1999, 2003). The
finding that language understanding may activate, e.g., the motor cortex, is, to a simulation
semanticist, taken as evidence that language is inherently an embodied, experience-based system
of knowledge. Those neurons that are utilized in the action of physical motion are also recruited
for the linguistic representation of said physical motion. Put differently, simulation semantics
eschews traditional abstract, amodal forms of representation (like, e.g., binary feature lists) in
favor of what Barsalou (e.g., 1999, 2003) terms “perceptual symbols”—subsets of experience
stored in memory to be retrieved and used later in cognitive domains like language and thought.
These perceptual symbols consist of complex patterns of neurological activation distributed
across varying neural networks. Thus, according to this model, the activation of a simple
linguistic concept may have ramifications in cognitive domains ranging from vision to touch,
gustation, audition, olfaction, and others (ibid.).

These patterns of language-induced co-activation are referred to as simulations, and
experimental evidence has converged to indicate that simulations are detectible phenomena. For
instance, in a seminal study during the field’s infancy, Glenberg and Kaschak (2002) documented
a phenomenon known as the action-sentence compatibility effect, where if the meaning of a recently
read (or recently heard) sentence is compatible with a physical action, that physical action is
affected (i.e., either facilitated or interfered with). To illustrate, if a subject reads a phrase like
*close the drawer* immediately before performing a congruent action (like, say, pressing a button by
moving their arm outward from their body) they may be faster to do so than if they had not seen
that previous sentence (or had seen a sentence bearing no overt semantic relationship with the
action performed).

Similar priming effects have been observed in other sensory modalities, as well. For
example, Yaxley and Zwaan (2007) found evidence of comparable effects occurring in the
domain of vision. In their experiment, they found that reading sentences like *Through the fogged*
goggles, the skier could hardly identify the moose caused experiment participants to respond more quickly to a blurred (or foggy) picture than a sentence like *Through the clean goggles, the skier could easily identify the moose*. In this same vein, Winter and Bergen (2012) showed that sentences describing scenes taking place at close and far distance primed experiment participants to identify large or small pictures (respectively). Furthermore, this effect was replicated in the auditory modality, where sentences describing scenes taking place at close and far distance primed experiment participants to identify sounds played at loud and quiet volume (respectively).

Although many of its studies are rooted in finding facilitation effects, simulation semantics has also found evidence of interference effects, especially within the visual modality. One key study that demonstrated such an effect was performed by Richardson and colleagues (2003). In this study, experiment participants were presented with sentences depicting horizontal scenes and vertical scenes (e.g., *The salesman points at the car* versus *The balloon floats through the cloud*). Following sentence presentation, participants were asked to identify images positioned along either the horizontal axis or the vertical axis of a computer screen. In this case, it was found that sentences with meaning congruent to the orientation of the subsequent image caused participants to be slower in completing the task. This surprising result, it was hypothesized, was due to the timing and sequencing of the task that subjects were asked to perform: Richardson and colleagues suggested that the neural resources used in perception are actually recruited for simulation-based representation, and are thus not as readily available for perception if simulation happens to be taking place.

In short, the findings of simulation semantics do indeed support a picture of language in which its cognitive representation is distributed across distinct neural networks traditionally thought to be implicated only in perception or action and not representation. Of course,
however, this finding comes with several caveats, some of them problematic, which will be discussed in the following section.

2.1.1 Problems with Simulation-Semantic Representation

There are four major problems currently facing the field of simulation semantics. They are (1) inconsistency of results across different studies, (2) variable interpretation of results across different studies, (3) inconsistency of results within individual studies, and, perhaps most importantly, (4) the inability of the model to account for abstract meaning. Each of these four problems will be treated below.

Firstly, the fact that both facilitation and interference effects have been observed in studies of the type described above poses a few different issues for simulation semantics. On the surface, the fact that both facilitation and interference effects are construed as simulation effects appears problematic, but the issue is, in truth, more complex than that. According to simulation semantics, the difference between registering a facilitation or interference effect is a matter of timing (see, e.g., Borreggine & Kaschak, 2006). Immediately after perceiving some form of linguistic input, the brain recruits perceptual resources (e.g., parts of the visual or motor cortex) for representation of the semantic content of said input. During this period, interference effects are observed. After the recruitment phase has ended, however, those same neurological regions are, for a brief period, primed for activation. During this phase, facilitation effects are observed. While this general narrative isn’t problematic per se, the devil is in the details. Figure 2.1 serves to illustrate this point:
**Figure 2.1:** Variation in reported response times (and effects) in simulation semantics studies. Effect times reported for Borreggine and Kaschak (2006) and Madden and Therriault (2009) are for Experiment One of both studies.

In Figure 2.1, we see the range of reported averages in nine different simulation semantics experiments across two different modalities (visual and motor simulation) and several different experimental methodologies. Items highlighted in green were described by their authors as facilitation effects, while those highlighted in red were described as interference effects. The figure illustrates (at least in part) several important points, which will be explicated in the following paragraphs.

First, looking at the above studies dealing with visual simulation, it becomes clear that developing a narrative of the specific time course of how such simulations play out would be a difficult undertaking. For example, in Madden and Therriault’s (2009) study of visual simulation effects, we see average response times in the 300-400 ms range (specifically, these numbers were
reported in experiment one of their study for postverbal reading times). However, other studies (i.e., Richardson et al, 2003; Bergen et al, 2007) show interference effects as taking place after the 300-400 ms window suggested in Madden and Therriault (2009). Given this, it immediately becomes apparent that such findings conflict with the general narrative of how simulations are supposed to work, as interference should, in theory, precede facilitation rather than follow it.

We also see, however, that other studies have produced visual facilitation effects that do fit with the interference-then-facilitation narrative, if we discount Madden and Therriault’s findings. These studies include Yaxley and Zwaan (2007) and Winter and Bergen (2012), who looked at visual features like blurriness and size (respectively) in visual simulations.

On the motor side, a wide range of response times can be observed across the studies featured above (Zwaan & Taylor, 2006; Bergen & Wheeler, 2010; Tseng & Bergen, 2005; Borreggine & Kaschak, 2006 [Experiment 1]). Of the four motor studies featured above, three of them appear to cluster roughly between 300 and 500 ms, which may indeed suggest a general trend. However, several other studies (including Borreggine & Kaschak (2006)) have listed other response times that diverge from the 300-500 ms window, such that they wouldn’t even fit on the above chart. For example, Glenberg and Kaschak’s (2002) landmark motor simulation study reported response times between roughly 1,650 and 1,850 ms.

Now, it needs to be emphasized that divergences in study results highlighted here are, to some extent, a product of methodological differences. For example, Borreggine and Kaschak’s (2006) response times also included reading time, whereas the other studies featured in the figure above did not. Furthermore, some of the visual studies featured above used a timed shape identification task while others used a timed picture verification task. The size and configuration of the button rigs used in some of the above studies varied, as well. There can be almost no doubt that these methodological differences have played some part in the inconsistency between
these results, but, in a way, that serves to bolster the point: no clear picture of the time course of simulations has emerged, and the methodological disparities between experiments do not serve to help that issue.

Another major issue facing simulation semantics has to do with authors’ variegated construals of significant findings. For instance, looking at Figure 2.1, one might notice that all of the motor effects featured are construed as facilitation. That would seem to imply that interference effects in motor simulation have yet to be observed, but, again, in reality, the picture is much more complex than that. For one, all of the motor studies highlighted above are based, at least in part, off of Glenberg and Kaschak’s (2002) landmark study, which famously demonstrated a simulation effect known as the “Action-sentence Compatibility Effect” (or ACE) whereby sentences with meaning congruent with the direction of a button press (either toward or away from the body) would cause people to execute said button press more quickly. Even though Glenberg and Kaschak’s study documents an effect that may be construed as either facilitation or interference (or a combination of both), because they terminologically framed the simulation effect that they found in terms of compatibility, a large number of subsequent studies within the same experimental paradigm have implicitly taken the stance that linguistic information primes motor movement, but does not interfere with it.

The problematic nature of this unidirectional approach to simulation reveals itself once the data itself is examined. For instance, a quick look at Glenberg and Kaschak’s results shows that their construal of the ACE as a simple one-way effect (be it facilitation or interference) may be oversimplifying things to a large extent. This is illustrated in Figure 2.2, below:
In Figure 2.2 (especially in the concrete transfer condition), facilitation, interference, or a combination thereof could all be posited as potential drivers of the effects observed. In the concrete transfer condition above, for example, there is a marked difference between away and toward responses in the yes-is-near condition, but not in the yes-is-far condition, as both the away and toward groups converge under these circumstances. In this case, one might be tempted to suggest that there are no simulation effects in the yes-is-far condition, and that in the yes-is-near condition, away sentences interfere with button press times while toward sentences facilitate them. However, if we look to the imperative condition above, we see that away button presses, in general, appear to take more time than close button presses, in which case we might construe the concrete transfer condition as one in which there are no observable effects in the yes-is-near condition, while we see away sentences generating facilitation effects and toward sentences generating interference effects in the yes-is-far condition. In short, without the inclusion of a control group, a baseline cannot be established. Because of this, the data (while statistically significant) lack a reference point to be compared against, and may thus be freely described as whatever effect best fits the narrative that the authors wish to present.
Similar results are observed in other simulation studies, as well. In Tseng and Bergen’s (2005) study of motor simulation effects in users of American Sign Language, we see a statistically significant result in essentially the same configuration as Glenberg and Kaschak’s finding listed above:

\textbf{Figure 2.3:} Left, another result that may be construed as either facilitation, interference, or a combination of both (reprinted from Tseng & Bergen, 2005)

In Figure 2.3, although the result in the \textit{Sem and Met signs} condition (in which Sign Language users provided semantic and metaphorical hand signs) is less symmetrical than that seen in Figure 2.2, it is still construable as facilitation, interference, or a combination of both. In this case, the authors construed it as a facilitation effect.

Unlike simulation motor studies, results in the visual modality have more stable interpretations. For example, in Bergen’s (2005) study dealing with literal versus metaphorical language and simulation effects in the upper and lower part of the visual field, we see a clear trend (in concrete sentences) illustrating an interference effect:
In Figure 2.4, we once again see a two-way interaction determining participants’ reaction times, this time taking place between noun meaning and location of object presentation. In this case, however, it is almost impossible to view the results as anything other than an interference effect, as the notion that a sentence with completely incongruent meaning could actually facilitate object classification seems rather implausible. In other words, the notion that a sentence describing action taking place in a given of the visual field would interfere with object categorization in that same part of the visual field makes sense; however, the notion that that same sentence would \textit{facilitate} object categorization in a different part of the visual field does not. Again, though, without a dedicated control group, establishing a baseline of how long users take to give responses in the absence of simulation effects, it is technically impossible to make this assessment, but, generally, this appears to be on more solid footing than similar assessments made in the motor simulation paradigm discussed above.

An additional problem facing simulation semantics currently is the problem of result consistency within individual studies. Often, and particularly within the motor simulation paradigm, significant results are obtained, but only in piecemeal fashion. Take, for example, the
results of Bergen and Wheeler’s (2010) study looking at the aspectual modulation of simulation effects. In this study, Bergen and Wheeler looked at the effects of both nouns and verbs on motor facilitation/interference in English as modulated by aspect (progressive versus perfect). According to the study, perfect aspect generated no significant results, whereas progressive aspect did. Crucially, the patterning of the results within the progressive aspect condition did not unfold as one would predict. This is shown in Figure 2.5:

**Figure 2.5:** One example of inconsistency in motor simulation results (reprinted from Bergen & Wheeler, 2010)

![Figure 2.5](image)

In Figure 2.5 above, we can see that noun-manipulated stimulus sentences (i.e., pairs of sentences that differ in the directionality implied by their direct object nouns, e.g., *The man touched the door* vs. *The man touched his nose*) generate a predictable pattern: In the *away* condition, participants hit the away button more quickly (and/or hit the toward button more slowly), while in the *toward* condition, participants hit the toward button more quickly (and/or hit the away button more slowly). However, in the verb-manipulated condition, a similar pattern is not observed. Instead, we see that no significant effect is generated in either direction in the *away* condition, while a very
robust effect is generated in the *toward* condition. Given the average times of the other results, it is likely that this constitutes a strong example of an interference effect, but, again, it is harder to make that judgment without simulation-free controls present.

Furthermore, consider again the results of Glenberg and Kaschak (2002). In Figure 2.2 (above), it should also be noted that in the concrete transfer condition, an effect distinguishing *away* and *toward* sentences is only observed in the *toward* condition. However, in the abstract transfer condition, we see that *away* and *toward* sentences are only observed to significantly differ in the *away* condition. Again, although these effects are indeed significant, it is unclear as to why they are not seen in both the *toward* and *away* conditions in both the abstract and concrete transfer situations.

The final and perhaps most difficult problem facing simulation semantics today is the issue of how it handles abstract meaning representation. Although the studies referenced above were cited for their issues with consistency in their results (and how their results were construed), what they have shown, at the very least, is that the general narrative of simulation semantics (wherein ‘non-linguistic’ regions of the brain are co-activated for representation) may indeed be true. In other words, the retrieval of perceptual states for representation, as outlined by Barsalou (1999) does appear to be supported by the variegated motor and visual facilitation effects mentioned above, even if it is somewhat difficult to neatly organize these findings. However, when simulation semantics attempts to go beyond the realm of the concrete and into the realm of the abstract, this narrative breaks down to a certain extent. This is what Zwaan (2014) terms the *secondary scaling problem* (p. 230):

Even for understanding a simple phrase like ‘a thief having to pay for stolen goods’, a ‘tower of abstraction’ has to be erected. This tower rests on concepts such as ownership of property, theft, social compulsion, and payment via fines or imprisonment, the latter concepts resting on notions of money and freedom. The secondary scaling problem is that this tower of
abstraction cannot be surmounted by systems that are equipped with grounded concepts only.

Even though Barsalou (1999) firmly states that abstract meaning can be handled via a perceptually grounded system of representation, results from experiments attempting to prove this have been rather inconsistent. For example, in a 2007 study, Bergen and colleagues found that sentences dealing with abstract motion like the stock price rose failed to elicit visual simulation effects from experiment participants. Similarly, metaphorical sentences with positive and negative valence like the policy failed also failed to elicit any sort of visual simulation effect.

Still, other studies within simulation semantics have successfully documented representation of abstract meaning via mental simulation. For example, Matlock, Ramscar, and Boroditsky (2005) found that sentences utilizing fictive motion (e.g., the road runs through the valley) affected participants’ reasoning about time. Similarly, Glenberg and Kaschak (2002), in the first-ever study documenting motor facilitation and interference effects, found a significant motor facilitation and/or interference effect for abstract instances of transfer (e.g., Sue told Liz the story).

Given the results seen above, it is difficult to conclude that simulation semantics, as a system, has been able to sufficiently account for how abstract meaning is represented. Here, however, it is suggested that a pluralist view (see, e.g., Dale, Dietrich, & Chemero, 2009; Dove, 2009; Zwaan, 2014) integrating simulation-semantic embodied representations with a more complete, experience-based system of cognitive and semantic representation (e.g., Fillmore’s frame semantics (see, e.g., Fillmore, 1976; Fillmore & Baker, 2011)) provides simulation semantics with a more solid foundation on which to represent abstract meaning, wherein representations are still experientially-grounded, but embodied simulations aren’t required for abstract meaning representation.
To summarize the above section, the field of simulation semantics has generated some groundbreaking and significant findings, but these findings have thus far failed to coalesce into a consistent larger picture, especially within the motor simulation paradigm. Across different studies, we see methodological inconsistencies, inconsistencies in timing, inconsistencies in how results are construed, and inconsistencies in whether or not abstract meaning can felicitously be represented via simulation. Further, within individual studies, we see inconsistencies in which results emerge across different conditions. In short, the findings of simulation semantics are, at this juncture, worth reporting, but not yet supported by a strong enough framework that can account for some of the inconsistencies seen in the data. The result is that the findings of simulation semantics are regarded by some as mere academic parlor tricks when they might not be viewed as such were they built upon a more solid theoretical foundation. The main theory presented in this dissertation, sub-event monitoring, provides a framework through which many of these issues can be ameliorated. Although it does not serve as a means of fully resolving all of the problems highlighted above, it will ultimately be shown that an approach to simulation that is better informed by lexico-pragmatic factors is more robust in the face of the problems highlighted in this section.

In the following section (Section 2.2), we will see that the lexico-pragmatic features that affect sub-event monitoring (e.g., usage frequency, lexical class, temporal density, etc.) are heavily involved in many aspects of language, including, ultimately, how it is represented. Again, if sub-event monitoring does indeed play a part in determining the character of mental simulations, it may therefore be possible that accounting for this in simulation semantics experiments can help to explain some of the variability reported above, as the aforementioned inconsistencies within and between experiments may (at least in part) be due to a lack of having controlled for these factors.
2.2 Lexico-Pragmatic Factors and Sub-Event Monitoring

The following sections outline the various lexico-pragmatic factors whose effects on mental simulation will be investigated in this dissertation. These factors include usage frequency (2.2.1), lexical class (2.2.2), temporal density (2.2.3) and others (2.2.4). As mentioned above, the theory of sub-event monitoring serves to link the above factors with observed simulation effects, helping to explain, at least in part, some of the variability that is often seen in simulation semantics results. Accordingly, in Section 2.2.5, sub-event monitoring will be fully explicated.

2.2.1 Usage Frequency

The primary feature that sub-event monitoring serves to link with mental simulation effects is frequency. As briefly mentioned in Chapter 1, some form of usage frequency can be found in each of the experiments featured in this dissertation. It is therefore central to the larger ideas, theories, and constructs of this work, and will be revisited with regularity.

Although simulation semantics and usage-based linguistics are considered to be separate streams of research, they have quite a bit in common, and do, in fact, overlap in some areas. Like simulation semantics, usage-based linguistics advocates for a view of a language faculty which is not divorced from other cognitive processes. Moreover, like simulation semantics, usage-based linguistics attests that linguistic patterning and (thus) representation are determined by trends across multifarious real-world discourse situations. The top-down view of Chomskyan parameter setting (see, e.g., Chomsky, 1981) is discarded in favor of a bottom-up picture of language acquisition in which linguistic categories and knowledge arise based on patterns of distribution in natural conversation.

To the usage-based linguist, the structure of human language is an emergent phenomenon (see, e.g., Hopper, 1987; Hopper, 1988; Bybee, 1998), and the engine that drives this emergence
is usage frequency (of both the type variety and the token variety). Frequency affects language in an abundance of ways. For example, in the short term (i.e., within the space of, say, a conversation), referential frequency affects how we shape discourse. Referents that are more frequently mentioned in the space of a conversation or topic sequence are encoded differently from referents that are only mentioned once or twice during that same timeframe (i.e., more frequently mentioned referents are more likely to be encoded nonlexically (e.g., as pronouns, clitics, zero anaphora, etc.)) (see, e.g., Cumming & Ono, 1997). This affects discourse participants’ recall of referents, occasionally creates instances of ambiguity, and, more generally, is heavily intertwined with the structuring of information in spoken conversation.

In the longer term, frequency has multifarious effects on, e.g., articulatory phonetics (see, e.g., Berkenfield, 2001), phonology (Bush, 2001; Bybee, 2001), semantics (see, e.g., Roberts, 2010 (Section 5)), morphology (e.g., Von Fintel, 1995), and syntax (e.g., Bybee & Thompson, 1997; Bybee, 2012) (for a more general review, see, e.g., Hopper & Traugott, 2003; Croft & Cruse, 2004; Bybee, 2007, 2010). Famously, English phrases like want to or going to have reduced to emergent modals wanna and gonna, respectively (see, e.g., Krug, 2001). This, according, to a usage-based linguist, is due to the high frequency of phrases like want to and going to. As they are more frequently used, the mind begins to store them not as separate words, but as single ‘chunks’ which may be more rapidly recalled and may also phonetically reduce across what are traditionally viewed to be word boundaries. The result is the creation of new lexemes that are amalgams of two or more previously independent words. Correspondingly, less frequent verbs which require a similar non-finite clause have not been observed to reduce in like fashion (e.g., love to study language does not reduce to luvda study language), so this is taken as evidence that frequency is the driving force behind such reductions.
We see similar frequency effects in other areas, as well. For instance, in a noteworthy study by Bybee (2001), French liaison was observed to more readily occur in more frequent phrases (e.g., cent années (one hundred years) is more likely to be liaised than cent escargots (one hundred snails)). Likewise, Bush (2001) notes that palatalization is more likely to occur in English across word boundaries in high-frequency collocations than it was in low-frequency ones (e.g., the /d/ in would you is more likely to palatalize to [dʒ] than the /d/ in supported you).

Frequency, while having the effect of reducing and streamlining some aspects of language, may also serve to entrench others. It is no secret that there appears to be a general trend across languages in which high-frequency verbs are more likely to be irregular, more likely to have withstood the force of analogical affixation. The irregular Spanish form, la mano, for example, remains a feminine noun because it is used so frequently and is so heavily entrenched in day-to-day language use that if one were to use the more paradigmatically consistent el mano, one would be viewed as having committed an egregious linguistic mistake. Similarly, the past tense of English go is based off of an entirely different verb than the present tense (wend) but the verb is used so frequently that it has somehow resisted language learners’ well-attested desire to replace went with goed (see, e.g., Bybee, 1985, Ch. 4).

In morphology, frequency seems to most often manifest itself in the area of grammaticalization. Heine and Kuteva (2007) define grammaticalization as “the development from lexical to grammatical forms, and from grammatical to even more grammatical forms” (p. 32), where, crucially, this reduction occurs within specific contexts (see also, e.g., Hopper & Traugott, 2003; cf. Norde, 2009). One example of this would be the development of the English suffix –ly, which comes from old Germanic *lik-, meaning body or same. This frequently occurring string’s meaning, pronunciation, and distribution have changed over time such that what used to be a word with a vastly different meaning now, for the most part, simply serves as an adverbial
suffix in modern English. Another classic example of grammaticalization is that of the English verb *will* (see, e.g., Wischer, 2008), which, as *willan* in old English, functioned as a lexical verb meaning *to wish* or *to want*. Over time, of course, that meaning has almost completely eroded, leaving primarily the future tense signification with which *will* is primarily associated in modern times.

According to the usage-based model, words which occur together frequently may be stored in the lexicon as ‘chunks,’ and these chunks may exhibit syntactic productivity patterns, giving rise to constructions (see, e.g., Goldberg, 1995, 2006; Michaelis & Lambrecht, 1996). Constructions, which emerge because of frequency, are the phrasal building blocks of language, and they may, as a result of their emergent nature, occasionally appear rather idiosyncratic. The syntax of constructions like *The Xer the Yer* (e.g., *The bigger they are, the harder they fall*) and the *way* construction (e.g., *Mary Lou Retton vaulted her way into America’s heart*) may seem nonstandard when compared with, say, a simple transitive clause, but nevertheless, they constitute productive linguistic patterns that have emerged because of frequency of use.

As demonstrated above, frequency can and does affect linguistic production. If, as simulation semantics claims (as laid out in, e.g., Barsalou (1999)), linguistic representation is the product of real-world experience, then it follows that linguistic representation must be built upon linguistic experience (at least in part). Thus, frequency of use must also affect representation to some extent. If, for example, the familiarity of well-rehearsed motor routines alters the pronunciation of certain words and phrasal chunks, then it follows that the act of using those words and phrasal chunks should alter and streamline their representation. What is proposed here is that when a linguistic representation becomes more streamlined due to high frequency, its pattern of motor activation (in pronunciation) isn’t the only thing to do so. In fact, the simulation evoked by said linguistic representation becomes more streamlined (but also more entrenched), as
Because of this streamlining, frequently used words and phrases are likely to be accessed more readily, therefore potentially skewing results in studies which rely upon response time as an outcome variable (see, e.g., Forster & Chambers, 1973).

This possibility raises several different interesting potential avenues for investigation, all of which have ramifications for theories of linguistic representation. It is unknown, for example, on what levels of linguistic representation (e.g., individual words, collocations, constructions) frequency can be demonstrated to affect the character of simulation. Furthermore, it is unknown if representations may temporarily be streamlined, as in the case of repeated reference over the space over a single discourse, or, equally, if they may be gradually and permanently streamlined over longer periods of time.

In a way, the question of the relationship between usage frequency and mental simulation is one that can tell us about the core of what mental simulation is and does cognitively. Although Barsalou (1999) famously claimed that mental simulations are used in a variety of different cognitive processes, subsequent research has treated them as a relatively unitary phenomenon. However, if, for example, the character of simulation changes depending on the familiarity of what is being simulated, then this does, in fact, suggest that mental simulations, not being an immutable form of representation, are much more epiphenomenal than previously believed. The latter possibility above is the position taken in this dissertation, and it will be revisited later in Chapter 7, where experiment results will be reviewed and synthesized.

2.2.2 Lexical Class

Linguistics has acknowledged lexical classes, in one form or another, for thousands of years. For example, the ancient Sanskrit grammarian Yāska identified four main lexical
categories in the fifth century, b.c.e., two of which corresponded to contemporary definitions of nouns and verbs (Matilal, 1990).

Although the litmus test for determining members of a given lexical class remains the same (find a set of words which behave in like fashion in several different linguistic environments), contemporary lexical classes are, of course, much more specific than the early parts of speech delimited by Yāska. To take one classic example, Vendler (1957) outlines four specific classes of verbs which he terms Aktionsart classes: activities, accomplishments, achievements, and states. In his work, Vendler notes that, for example, activity phrases readily fit with temporal questions like for how long, but achievements and states (of the permanent variety) do not:

(1) For how long did the child swim? (activity)
(2) #For how long did the train arrive? (achievement)
(3) #For how long was the bus driver Canadian? (state)

Furthermore, accomplishments may felicitously appear with terminative temporal questions beginning with how long did it take to, while activities and states may not:

(4) How long did it take the carpenter to build the chair? (accomplishment)
(5) #How long did the child to swim? (activity)
(6) #How long did it take the bus driver to be Canadian? (state)

Similarly, questions beginning with at what moment serve to differentiate achievements from other Aktionsart classes:

(7) At what moment did the train arrive? (achievement)
(8) #At what moment was the bus driver Canadian? (state)

The above Aktionsart classes outlined by Vendler were not generated ad hoc, but were instead posited based upon the distributional patterns of various verbs in natural language. Such distributional differences between words can reflect semantic differences between lexical classes which are the product of differences in tangible, embodied experience, as in the case of, say, count nouns and mass nouns, or they can reflect more subtle semantic differences, as in the case
of Rappaport-Hovav and Levin’s (2008) verbs having only a caused possession meaning (e.g., give) and verbs having both caused motion and caused possession meanings (e.g., send), both of which participate in the so-called English dative ‘alternation.’

Rappaport-Hovav and Levin note, for example, that even though verbs like give and send are both ditransitive verbs which denote scenes wherein an entity is transferred, there are distributional differences between these verbs (and their ilk) which suggest that they are, in fact, members of different lexical classes. For instance, send appears to allow the act being described to have not been successfully completed, whereas give does not:

(9) I sent the gift to Lisa, but she never received it.
(10) #I gave the gift to Lisa, but she never received it.

In a similar fashion, one can ask destination-related questions about theme arguments with send-type verbs, but not with give-type verbs:

(11) Where did you send the gift?
(12) #Where did you give the gift?

In short, although at first give and send-type verbs appear to be nearly semantically identical, an analysis of their distributional patterns reveals that they are likely members of two (or more) distinct lexical classes with unique meanings and cognitive representations.

Simulation semantics studies often use ad hoc classes based on the methodology of the experiment being performed in combination with a norming study. For example, a motor simulation study may use verbs which denote manual motion toward the body along with verbs which denote manual motion away from the body. Those verbs which are most agreed upon as fitting their respective classes will be used for generating stimulus sentences. The issue, of course, is that these ad hoc classes disregard the independently motivated lexical classes of which their members might be a part.
The theory of sub-event monitoring suggests that lexical classes may, in fact, have a fairly significant role to play in shaping mental simulation, even if these classes are ostensibly not differentiable by any overtly ‘embodied’ semantic characteristics or features. This is because lexical classes are shaped by a combination of (a) world knowledge and (b) linguistic and extralinguistic experience. Because of this, lexical class distinctions have the power to evoke contrasts ranging from the tangible to the intangible, from the concrete to the abstract. For example, classes may differ in terms of, say, their temporal density (to be covered immediately below) or their markedness (to be covered in Section 2.2.4), both of which are lexico-pragmatic features implicated in the theory of sub-event monitoring. This means that certain lexical classes may naturally be inspected more closely than others during simulation. These differences in inspection, it will ultimately be shown, strongly affect mental simulation.

In this dissertation, both Vendler’s Aktionsart classes and Rappaport-Hovav and Levin’s caused possession only and caused motion and caused possession classes will be examined in experiments three and four (Chapters 5 and 6), respectively. It will ultimately be shown that lexical class does indeed significantly predict simulation effects in two different simulation modalities.

2.2.3 Temporal Density

Language and time are two thoroughly intertwined entities, and there is an impressive variety of means for expressing time that language has at its disposal. In English, time not only presents itself in grammatical areas like tense and aspect, but in semantic areas (e.g., Aktionsart class), as well. According to the Oxford English Corpus, the word time is even the most frequent noun in the English language that isn’t a pronoun. As astutely noted by many scholars of metaphor (see, e.g., Lakoff & Johnson, 1980), time is a thoroughly abstract phenomenon, and many aspects of language help to ground that abstraction in the physical reality around us. We
may, for example, speak of time as ‘dragging on’ or ‘flying by’ even though it is not a physical object that can fly or drag. We may be ‘coming up’ on an event even though we are sitting still in space, and we may ‘look back’ at our pasts without physically turning around. In short, time is fundamental to both embodied experience and language, and, as such, it presents itself as a logical avenue for inquiry in simulation semantics study.

Of particular interest for our purposes here is the notion of temporal density, which serves as a major factor in sub-event monitoring. Following the work of Mittwoch (1988), event structure information as instantiated in human language may be metaphorically conceptualized as a window (or representational space) looking over a timeline. Sometimes, entire bounded events may fit within this representational space, while at other times, due to the constraints of the space, we may not be able to see these events’ startpoints and/or endpoints. These events may have homogenous internal structure (as, for example, in the case of lexically stative verbs) or heterogeneous internal structure (as, for example, in the case of telic accomplishment predcations like build a chair) (see, e.g., Comrie, 1976; Bickel, 1997; Michaelis, 2011).

Following Michaelis (2011) (as well as Bickel (1997)), Aktionsart classes are decomposable into two components: \( \varphi \), which symbolizes a phase, and \( \tau \), which symbolizes a transition or boundary. Lexically stative verbs may simply be described as simple phases \( ([\varphi]) \), whereas telic verbs like accomplishments necessarily involve a buildup to some sort of completion and are thus best described as a complex combination of phases and transitions \( ([\tau \varphi \tau \varphi]) \). A punctual verb, like, say, release, may simply be described as \( [\tau \varphi] \). Crucially, these Aktionsart representations can and do interact with grammatical tense and aspect. For example, the use of the past progressive with a verb of accomplishment (e.g., Paul was building a chair) selects for the phase portion of its Aktionsart operator representation \( ([\tau \varphi \tau \varphi]) \). Conversely, for those verbs that lack a phase
portion (i.e., semelfactives) (and are thus composed of only a transition/boundary), stativization requires that this transition be iterated (e.g., *The light was flashing*), resulting in multiple transitions within the representational space of such predications.

With the above in mind, temporal density can be defined as the following:

(13) The number of transitions within the representational space of a predication

The following figure (Figure 2.6) exemplifies this notion:

*Figure 2.6: Transitions (or lack thereof) in the representational space of three different predications.*

Above we see diagrammed the representational spaces of three different predications, each of which has a different internal structure. The first, *The woman was Swedish*, exemplifies a homogenous state that neither starts nor ends in the representational space. The second, *The woman blinked*, represents a punctual single event that both starts and ends in the representational space. The third, *The woman was blinking*, represents a punctual event that has been stativized (via the progressive) and is therefore given an iterative reading.
Crucially, temporal density can be altered by a number of different factors, the first being speed of event succession. This is exemplified in Figure 2.7, below:

**Figure 2.7:** The hypothetical effect of a manner adverb like quickly on temporal density.

In Figure 2.7, we can see that the latter sentence, *The woman was blinking quickly*, has comparatively more transitions in its representational space by virtue of the fact that the adverb *quickly* has been added, denoting a marked increase in iteration. Thus, of the above two sentences, the latter has a greater temporal density than the former.

Secondarily, although Mittwoch’s conceptualization of the representational space window is fixed in size, that shouldn’t necessarily be the case, as linguistic content is well-documented to vary in terms of its temporal scope. Consider the following (Figure 2.8):
Above, we see an illustration of differences in temporal scope ultimately leading to differences in temporal density. Specifically, if we assume a fixed transition rate across instances of the same verbal predication (save for any modifying adjuncts which would change the size of our representational space), then the number of transitions within a given representational space should change as temporal scope changes. Specifically, the latter sentence has a greater number of transitions in its representational space than the former.

One well-attested facet of the human experience of time is its variability. As mentioned earlier, we may refer to time as ‘dragging on’ or ‘flying by,’ depending on the circumstances. These types of experiences fall under the umbrellas of temporal dilation and temporal compression, respectively. According to Evans (2003) (summarizing Flaherty (1999)), the circumstances that give rise to temporal dilation and compression are, in fact, predictable. Of temporal dilation, Evans writes (p. 20):

"Experiences which give rise to a higher density of information processing, and hence in which time appears to pass more slowly (protracted duration), include suffering and intense emotions, violence and danger, waiting and boredom, concentration and meditation, and shock and novelty. As the subject is consciously attending to the stimulus array, a greater density of information processing occurs. Given that our experience of duration appears to correlate with the amount of memory taken up ... then if more of the stimulus array is attended to, more memory is required to store and process what is being attended to, and consequently it is to be expected that we should actually experience the duration as being more protracted, which is what we find."

And, of temporal compression (ibid.):
Experiences which produce a lower density of information processing, and hence in which time appears to ‘pass more quickly’ (temporal compression), include those which involve routine complexity … Habitual conduct results in little of the stimulus array being attended to, resulting in low density of information processing. Accordingly, time seems to have passed ‘quickly’.

Given the above, we should expect to see effects of temporal dilation in linguistic predications in which the representational space delineated above depicts either the intensely novel (i.e., scenes with highly infrequent words, a lack of internal homogeneity, multiple transitions) or the intensely sparse (i.e., scenes with very little information/very few transitions). Conversely, in those instances where information in the representational space is complex but predictable (e.g., predications with heterogeneous but familiar event representations), we expect to see temporal compression effects.

In this dissertation, a modified version of the above approach to subjective temporal variability is adopted. Namely, in keeping with pacemaker-accumulator models of neural timekeeping (see, e.g., Gibbon, Church, and Meck, 1984), I predict that predication types which are more closely inspected in sub-event monitoring (i.e., those predications which depict scenes that are more temporally dense, less expected, more dynamic, etc.) will generate temporal compression effects, while predication types which are less closely inspected (i.e., those predications which depict scenes that are less temporally dense, more expected, less dynamic, etc.) will generate temporal dilation effects. In accordance with temporal cognition models advocated for by Droit-Volet and Meck (2007) (as well as Droit-Volet and Gil (2009)), it is postulated that high-inspection simulation events will divert resources that are otherwise devoted to neural timing, therefore yielding temporal compression effects. Correspondingly, it is postulated that low-inspection simulation events will cause the opposite effect: a lack of inspection in simulation affords additional resources to neural timekeeping that may not always be present, effectively slowing down the rate at which time appears to be passing. This model will be
revisited throughout the dissertation and built upon in Section 5.5.5 in light of the results of experiments two and three.

2.2.4 Other Lexico-Pragmatic Factors

Thus far, three different lexico-pragmatic factors have been introduced as being features that sub-event monitoring links with embodied mental simulation. These are usage frequency, temporal density, and lexical class. As one might imagine, however, there are a wide variety of lexico-pragmatic features, and many (if not all) of them may have an effect on how simulation plays out. For instance, in experiment three (Chapter 5), a subset of the critical stimuli presented are dedicated to testing how canonical ‘expectedness’ affects temporal simulation. In this case, stimulus sentences are controlled for usage frequency (in its various forms), but some sentences are more ‘prototypical’ than others (e.g., the messenger rode the bicycle versus the messenger rode the dragon). Additionally, in experiment four (Chapter 6), constructional markedness (see, e.g., Andersen, 1989) is demonstrated to significantly factor into motor simulation, as one constructional member of the English dative opposition (the ditransitive construction) is marked for change of possession, while the other (the oblique goal construction) is not, leading to differences in inspection in sub-event monitoring.

2.2.5 Sub-Event Monitoring Explained

Simulation semantics, as a whole, has shown a tremendous amount of potential, but a lack of consistency across and within studies has hampered this to some extent. However, taking into account the notion of sub-event monitoring may help to explain some of this variability and, ultimately, bridge the gap between usage-based linguistics and simulation semantics. Sub-event monitoring can be defined as the following:
The degree to which the salient simulation-based features of a linguistic predication are inspected due to lexico-pragmatic factors.

As stated immediately above, sub-event monitoring is affected by several different lexico-pragmatic factors, but ones that figure in most pertinently are usage frequency, lexical class, and temporal density. The way that these factors can interact with one another is ultimately quite complex. This complex interaction can be broken up into three tiers, which are summarized below.

Before launching into a discussion of the components that govern sub-event monitoring, however, I will briefly define and summarize what sub-event monitoring is. Firstly, it needs to be pointed out again that simulations are not, according to Barsalou (1999), monolithic phenomena. In other words, a simulation is not a single neurological state that is consistently reactivated, but rather, it is a dynamic pattern of activation which changes according to the content being simulated. For example, when we simulate a sentence like you are throwing the baseball, it is not the case that the entire act of throwing a ball is activated simultaneously (i.e., clutching an object, readying to throw, starting to throw, releasing the object, and following through). Rather, simulations reflect the progressions of the real-world events off of which they are based. In the case of simulating the act of throwing a baseball, for example, we would expect the earlier stages in a typical sequence of throwing to be represented first, and the latter stages to be represented last.

It is therefore the case that simulations are composed of sub-events. However, it needn’t be the case that all of these sub-events be inspected and utilized in every instance of mental simulation. For instance, in simulating a highly frequent phrase like give money to X, it may not be necessary for us to inspect every sub-event of such a quotidian transaction. We would therefore expect such a phrase to be processed (and simulated) more quickly. However, when we simulate
a less common phrase, like *bequeath seashells to X*, such an unorthodox event may indeed be inspected more closely.

It may also be the case that certain aspects of the sub-event structure of a simulated event (e.g., its temporal density) cause us to monitor it more closely or less closely. As will be discussed in experiment three (Chapter 5), stative predications are viewed to have homogenous internal structures, while, say, accomplishment predications do not have homogenous internal structures. It thus behooves the simulator to monitor accomplishment (i.e., dense) predications more closely while doing the opposite for stative (i.e., non-dense) predications.

Having established the basic fundamentals of sub-event monitoring, we can now turn to the various features that it serves to link with mental simulation. As stated above, these features are organized into three different tiers, each of which contributes to sub-event monitoring in a unique way.

The first tier of sub-event monitoring consists of the basic semantic content to be simulated. For example, sentences dealing with the literal act of *throwing*, regardless of the lexico-pragmatic features they may contain, access a mental simulation that involves the motor or premotor cortex. As Barsalou (1999) points out, this content may have any combination of sensory features, such that a verb phrase like *taste salty* may primarily access, say, the gustatory cortex, but also secondarily access the olfactory cortex and manual motor cortex, as smell and manual action are both heavily involved in the act of consuming food, as well.

The second tier of sub-event monitoring content is usage frequency. This includes long-term usage frequency, short-term usage frequency, various other forms of token frequency, type frequency, and so forth. In general, high-frequency words and phrases are processed more quickly than low-frequency words and phrases (see, e.g., Forster & Chambers, 1973, Forster 1976; Balota & Chumbley, 1985; Ellis, 2002). Given this, frequency, in and of itself, can have
substantial effects on how simulation effects play out (or how they are construed). For example, a relatively frequent verb phrase like *throw the ball* is likely to be processed more quickly than a less frequent verb phrase like *throw the candlestick*. As such, one would surmise that simulation-based facilitation and (subsequent) interference effects may arrive more quickly for the former phrase than for the latter. This, of course, has the potential to confound results, as the average response time for a frequent phrase in an incongruent condition could potentially overlap with the average response time for an infrequent phrase in a congruent condition.

Finally, the third tier consists of other lexico-pragmatic features that may interact with the second tier. These include lexical class, temporal density, markedness, and others. Specifically, these features interact with the second tier by *diminishing* or *enhancing* frequency effects. Specifically, predication types with third-tier features which require a high degree of monitoring are associated with *diminished* frequency effects, and predication types with third-tier features which require a low degree of monitoring are associated with *enhanced* frequency effects. Thus, for example, if we were looking at a high-temporal-density (and thus, high monitoring) verb class and a low-temporal density (and thus, low monitoring) verb class, and plotted the simulation effects of all of the members of that verb class by the frequency of each individual member, the frequency-based modulation of simulation effects would occur over a more extreme range in the case of the low-temporal-density (i.e., low-monitoring) class and over a smaller range in the case of the high-temporal-density (i.e., high-monitoring) class. This relationship is illustrated in the abstract in the following figure (Figure 2.9):
Looking at the above figure, it’s clear that different levels of sub-event monitoring could strongly affect how simulation data are interpreted. Specifically, as the figure illustrates, the configuration of average participant responses for high-density and low-density predications has the potential to reverse as frequency increases. This, of course, has major ramifications how simulation results are construed and interpreted. Experiments whose stimuli are not controlled for frequency in some form are therefore essentially not reporting the entire picture.

To summarize the various components of sub-event monitoring: (1) some sort of basic semantic content is represented via mental simulation, (2) the frequency of said content affects the amount of time it takes to process said simulation, where less frequent content is processed more slowly, and more frequent content is processed more quickly, and (3) other lexico-
pragmatic factors enhance or diminish the effects of frequency mentioned above. This general scheme is illustrated in the following figure (Figure 2.10):  

**Figure 2.10:** A diagram illustrating the basic relationship between the different tiers of sub-event monitoring.

In Chapter 7, it will be shown that the above model is able to account for the distribution of the vast majority of the data seen in this dissertation, and a thorough discussion regarding its strengths and shortcomings will be provided.

Finally, a quick recap of the sub-event monitoring features mentioned above is given in the table below (Table 2.1). In it, sub-event monitoring features are organized both by tier and the experiment in which they appear.
Table 2.1: A table showing the lexico-pragmatic factors implicated in sub-event monitoring to be featured in this dissertation.

<table>
<thead>
<tr>
<th>Feature</th>
<th>Tier</th>
<th>Experiment One</th>
<th>Experiment Two</th>
<th>Experiment Three</th>
<th>Experiment Four</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simulation Content</td>
<td>One</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Usage Frequency</td>
<td>Two</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Temporal Density</td>
<td>Three</td>
<td></td>
<td>✔</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Expectedness</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lexical Class</td>
<td></td>
<td></td>
<td></td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Markedness</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✔</td>
</tr>
</tbody>
</table>

Having outlined (a) the major theoretical issues at stake, and (b) the linguistic features that are central to our investigation of these issues, we can now turn to the two experimental methodologies that will be used in this dissertation: the motor facilitation and interference paradigm and the temporal estimation paradigm. Summaries of both follow immediately below.

2.3 Methodologies

This section outlines the two major methodologies to be used in this dissertation. The first, the motor facilitation and interference paradigm, is well established within the simulation semantics literature. The second, the temporal estimation paradigm, constitutes a new methodology to be used in simulation semantics. Explications of both follow below.

2.3.1 The Motor Facilitation and Interference Paradigm

The motor facilitation and interference paradigm to be used in experiments one and four dates back to a landmark (2002) study performed by Glenberg and Kaschak. Since the study’s publication, numerous other studies have used the same methodology (or variations thereof) and,
as discussed above in Section 2.1.1, achieved noteworthy but varying results (for a general review, see, e.g., Fischer & Zwaan, 2008; Anderson & Spivey, 2009). As in other simulation semantics experimental methodologies, at the core of the paradigm lies the idea that representation utilizes some of the same resources as execution (and/or perception) (Barsalou, 1999), and, therefore, that activating certain linguistic representations should, in some form, affect motor execution. This notion has been supported (if not outright verified) by various neuroimaging studies (see, e.g., Pulvermüller, 1999, 2001; Hauk, Johnsrude, & Pulvermüller, 2004) which have demonstrated that the act of reading words that deal with concrete motor actions and concepts activates the motor cortex, and that the act of reading words dealing with visual scenes similarly activates the visual cortex. Hauk and colleagues (2004) even showed that verbs denoting physical actions using different body parts (i.e., hand, leg, face) activated the specific sub-regions of the motor cortex responsible for controlling those body parts. These findings paint a picture of lexical representation that is highly distributed across multiple neural regions but is nevertheless predictable. For example, a verb’s phonological representation may be located in one region, its selection restrictions in another, and its motor representation (if any) in another (i.e., the motor cortex).

Like the studies mentioned directly above, the motor facilitation and interference paradigm is also able to detect language-induced motor-cortical activation. The extent of this activation is gauged via arm movement time in a button press task (to be discussed in greater detail below). Given this means of observation, this methodology has considerably less granularity than, say, an fMRI study, but it is also good for collecting large amounts of data for users’ reactions to any number of different linguistic items, classes, and configurations. Beyond that, fMRI equipment is often forbiddingly expensive and relatively complex, whereas motor facilitation and interference studies are markedly easier to put together and implement.
In the motor facilitation and interference paradigm, participants are presented with some form of linguistic stimulus (either written or auditory) and then given a binary response choice, one of which is physically congruent in some way with the linguistic stimulus, and the other of which is physically incongruent with the linguistic stimulus. In the majority of motor facilitation/interference studies, this is done through a rig consisting of three buttons (often a keyboard turned ninety degrees from its typical orientation on the transverse (i.e., horizontal) plane). This is illustrated in Figure 2.11 below:

*Figure 2.11: A photograph of a typical motor facilitation/interference experimental setup.*

In the three-button rig illustrated above, we see a central yellow button that is used to display stimulus sentences. Stimulus sentences are displayed only as long as the yellow button is held down, and upon its release, these sentences disappear. After reading a stimulus sentence, participants may then choose either the green or red button. Crucially, these buttons are oriented such that pressing one button requires arm movement away from the body while pressing the other button requires arm movement toward the body. Thus, under the right experimental conditions, investigators may gather both motion-congruent and motion-
incongruent responses by providing stimulus sentences that are compatible or incompatible with
the location of the green button. For instance, a sentence like *I am throwing the ball* would be
congruent with an away button press but incongruent with a toward button press. Conversely, a
sentence like *I am touching my nose* would be congruent with a toward button press but incongruent
with an away one.

As stated earlier in Section 2.1, it is possible to vary this paradigm in a number of
different ways. The original version of the paradigm used by Glenberg & Kaschak (2002), for
example, implemented the basic setup mentioned above, but included reading times in their
button press data. Other studies (e.g., Bergen & Wheeler, 2010) have since abandoned the
inclusion of reading times in button press data and instead have opted to have the arm
movement phase happen discretely after the reading phase.

Another version of the motor facilitation and interference paradigm abandons the three-
button rig in favor of a dial that may be turned in a clockwise or counterclockwise motion (see
e.g., Zwaan & Taylor, 2006). In this version of the paradigm, the individual words of a sentence
are presented on a screen, and the subject is instructed to turn the dial in order to cause the
words of the sentence to progress. The time it takes for the subject to move the dial is recorded
for each individual word, and, depending on the congruency (or lack thereof) of the semantic
content of the certain words (Zwaan and Taylor looked primarily at verbs), this time will vary. If,
for example, subjects are asked to turn the dial clockwise in order to advance the words within a
sentence, the act of reading a word that denotes an incongruent (i.e., counterclockwise) rotation
(e.g., *unscrew*) should affect subjects differently than the act of reading a word that denotes a
congruent clockwise rotation.

Within the button press version of the motor facilitation and interference paradigm, there
are several other ways in which the study setup itself can vary. For instance, numerous button
configurations may be used, and, as mentioned earlier, the rigs on which these configurations are built frequently vary in terms of their size and dimensions. The timing of stimulus presentation is another factor that varies across studies, and, as astutely pointed out by Borreggine and Kaschak (2006), it is a crucial one in observing motor facilitation and/or interference effects. In their study, Borreggine and Kaschak looked at the effects of response cue timing on facilitation and interference effects. After stimulus sentence presentation, participants were asked to wait until they were given a cue before they could initiate their button presses. The cues were timed at 0 ms, 50 ms, 500 ms, and 1000 ms, and results varied dramatically:

**Figure 2.12:** The difference between incongruent and congruent response times after four response cue time intervals (Experiment 1=0ms, Experiment 2=50 ms, Experiment 3=500 ms, Experiment 4=1000 ms) (reprinted from Borreggine and Kaschak (2006))

In the above figure, we see that at a cue time of 0 ms, motor simulations produce strong differences between incongruent and congruent conditions, where the congruent condition requires less execution time than the incongruent condition. At 50 ms, however, this effect diminishes greatly and nearly disappears. At 500 and 1,000 ms, the effect actually reverses, wherein we observe the incongruent condition as generally taking less time than the congruent condition. In short, in the motor facilitation and interference paradigm, timing is absolutely crucial to whether results are observed and how those results are construed.
Overall, the merits (or lack thereof) of the motor simulation and interference paradigm are a matter of perspective. Given the issues with consistency between studies that are presented above, it would be easy to simply discard it in favor of an alternative methodology. However, we would be remiss to do so. If these studies begin to adhere to a more consistent methodological model, a much clearer picture of simulation motor effects may begin to emerge. Further, the number of contributing factors that are as of yet unaccounted for in motor simulation is far too great for one to ignore the potential for significant findings that may help to explain some of the variation in the data.

While it lacks the precision of, say, fMRI imaging, the motor facilitation and interference paradigm allows us to test the simulation effects induced by any number of different linguistic phenomena, be they individual lexemes, entire lexical classes, grammatical patterns, alternating constructions, different discourse sequences, etc. And it demonstrates a remarkable finding: linguistic knowledge—once considered to be a wholly abstract apparatus sequestered to a specific module of the brain—actually manifests itself in a distributed, embodied fashion.

### 2.3.2 The Temporal Estimation Paradigm

The second experimental paradigm to be used in this dissertation constitutes a new experimental methodology in the field of simulation semantics. This methodology, termed *temporal estimation*, aims to show not only that mental simulations have a temporal component, but also that this temporal component may change in character depending upon the presence or absence of certain lexico-pragmatic features (most notably, temporal density).

Although work within simulation semantics has focused heavily on more traditional sensory modalities (e.g., vision, audition, etc.) (see, e.g., Barsalou, 1999), little attention has been devoted to investigating the nature of the temporal component of mental simulations (however,
see Zwaan, 2008; Evans, 2008). This is, in all likelihood, due to the fact that the subjective experience of time is a notoriously abstract and difficult phenomenon (or, rather, set of phenomena) to quantify and measure in an experimental setting. Evidence from numerous sources has suggested that neural timekeeping is not done by one central internal clock, but through a collection of distinct neural regions, and these regions “work in concert, but are nonetheless separable” (Pariyadath & Eagleman, 2007, p. 1264) (for a review, see Evans (2003)). This leads to some findings about time perception that one might consider counterintuitive. For example, Pariyadath and Eagleman (2007) demonstrated that subjects’ judgments of duration could be manipulated by presenting them with an unexpected event (which caused the subjective slowing down of time, or temporal dilation). It was hypothesized, for instance, that a fixed tone played during every stimulus presentation would be perceived as being lower in pitch while an oddball stimulus was presented because this unexpected event would briefly slow down time for experiment participants, thus causing a subjective lowering in pitch frequency. However, during these unexpected events, other aspects of perception were found to be unaltered even though subjects did experience a temporal dilation effect (see Stetson, Fiesta, and Eagleman (2007) for a similar finding).

The above (and other) findings have painted a picture of a subjective time that is highly compatible with certain aspects of mental simulation. Those events (or sequences thereof) that are highly infrequent or unexpected cause the subjective slowing of time (i.e., temporal dilation), while those events that are routine may cause the subjective speeding up of time (i.e., temporal compression). Also, a complete lack of salient events (as experienced, e.g., during times of extreme boredom) may also cause temporal dilation (again, see Evans (2003), ch.1 for a review). In this regard, temporal variability can be viewed as an embodied phenomenon, and if certain
events or classes of events are associated with temporal compression or dilation, it logically follows that this variability should be encoded in simulation-based representation.

Given that the well-documented (but complex) variability in temporal perception is likely encoded, in some fashion, in simulation-based representation, it is hypothesized that certain elements of this variability could be observed in a simulation-semantics-based experimental setting. Specifically, as discussed in Section 2.2.3 above, according to sub-event monitoring, it is hypothesized that the sub-event structures of high-temporal-density predications will be closely inspected and, because this inspection diverts resources away from neural timekeeping, these predications will be associated with temporal compression effects. Conversely, it is hypothesized that low-temporal-density predications will be less closely inspected, resulting in the opposite effect (i.e., temporal dilation effects). These hypotheses will be revisited in experiments two and three (Chapters 4 and 5) in light of the findings generated by these studies.

It is clear that there are rich and multitudinous avenues for potential study using the temporal estimation paradigm. Temporal information presents itself in an astonishing variety of ways in human language, and if that temporal information could actually be proven to affect the speed of one’s ‘internal clock,’ that would constitute a noteworthy finding. What follows below is a brief explication of the methodology used in the temporal estimation paradigm and how it may prove that the language we use affects the time we perceive.

The temporal estimation paradigm is rooted in the simple idea of subjective time being a variable phenomenon. To restate a few claims made earlier in Section 2.2.3, it is common wisdom that subjective time, as we experience it, may travel at different speeds (e.g., the time flew by vs. the hours dragged on). If, according to the story of both simulation semantics and cognitive-functionalist linguistics, mental representation (and thus, mental simulation) is the product of real-world experience, then the sensation of varying subjective time should be encoded in some
fashion in mental simulation. Accordingly, the simple act of, say, reading a sentence in a book, may, under the right conditions, be enough to change the speed of one’s own internal clock, if only for an instant and in an imperceptible fashion.

The temporal estimation paradigm is able to gauge such imperceptible changes using the following sequence of events on a trial-by-trial basis: First, a simple auditory stimulus is presented (in this case, a sine wave tone of 440 Hz). Following that, linguistic stimulus is given (a phrase or sentence that participants must read aloud). Lastly, participants are asked to hold down a button which recreates the auditory stimulus (the 440 Hz sine wave tone) and reproduce its temporal length. Under conditions that do nothing to affect participants’ internal clocks, participants’ mean response time should be extremely close to the length of the tone initially presented. This process is illustrated in Figure 2.13 below:

**Figure 2.13:** A simple illustration of a single trial in the temporal estimation paradigm.

![](image)

In the above chart, we see the simple three-part structure of the temporal estimation paradigm. The tone presentation phase is listed as lasting two seconds, although other intervals can be used in order to avoid making the task too repetitive for participants. During the stimulus presentation and reading phase, participants may be either (a) asked to read stimulus sentences out loud, or (b) given visual or auditory stimulus sentences and follow-up questions in order to
ensure comprehension. Finally, during the tone reproduction phase, participants hold down a button for what they perceive to have been the duration of the tone given in the tone presentation phase.

Using the above sequence of events, if, for example, a participant hears a tone and then reads a sentence that changes the speed of their internal clock, when they are asked to reproduce that tone, the participant will think that they are being accurate (even when, say, counting using internal monologue) when, in reality, their estimation will be pushed in one direction or the other by the effects of sub-event monitoring. In those instances where participants are experiencing temporal compression (where, e.g., one subjective second is composed an objective time interval longer than one objective second), we expect them to press the tone reproduction button for more time, on average. In those instances where participants are experiencing temporal dilation (where, e.g., one subjective second is composed an objective time interval shorter than one objective second), we expect them to press the tone reproduction button for less time, on average.

It should be briefly noted here that the above construal of temporal simulation effects (i.e., the overestimation corresponds to temporal compression while underestimation corresponds to temporal dilation) represents, in some ways, a break from previous literature, as time perception literature is filled with contradictory claims about what factors may cause one’s temporal perception to change. As Hicks, Miller, and Kinsbourne (1976) point out, “much of the inconsistency [in temporal perception literature] is in different experimenters’ conclusions, and not their results. The variability in conclusions stems from a confusing use of terms (e.g., overestimation and underestimation of time) and a lack of appreciation of procedural differences.” (p. 719).

Hicks and colleagues suggest that a major source of this inconsistency is some confusion about prospective versus retrospective temporal estimation. Specifically, prospective and
retrospective temporal estimation famously pattern opposite of one another. For example, if a subject is asked to give a temporal estimation *while* experiencing, say, temporal compression effects, they will *overestimate* the amount of time that has passed. However, if that same subject, now no longer experiencing temporal compression, is asked to retroactively how much time had passed while they were experiencing temporal compression, they will *underestimate* the amount of time that passed.

To problematize things further, there are actually *two* different forms of retrospective temporal estimation that people may perform, and, again, they pattern opposite of one another. These are *baseline temporal state judging altered temporal state* and *altered temporal state judging baseline temporal state* (which the temporal estimation paradigm employs). If we define *baseline* temporal perception as one’s default cognitive state in which subjective time and objective time align, then *altered* temporal perception refers to those cognitive states in which subjective time and objective time *disalign*. In this case, when objective time appears to pass more slowly than subjective time, one is said to be experiencing temporal dilation (i.e., the sensation that time is dragging on). Correspondingly, when objective time appears to pass more quickly than subjective time, one is said to be experiencing temporal compression (i.e., the sensation that time is flying by).

To illustrate how disalignment between subjective time and objective time (or a lack thereof) could affect how temporal estimation results are construed, consider the following two hypothetical retrospective temporal estimation scenarios:

(1) **Baseline Judging Altered**: A participant is given a stimulus which causes a period of temporal compression that lasts five objective seconds. However, because they experienced temporal compression during this period, it felt like less than five subjective seconds. Following this period of temporal compression, when this participant, now in
his or her baseline state, is asked to report or reproduce the time interval that they experienced, they will provide a temporal interval that is less than five seconds. In this instance, then, underestimation indicates temporal compression effects, and overestimation indicates temporal dilation effects.

(2) **Altered Judging Baseline:** A participant experiences an interval of five subjective/objective seconds while in their baseline state. Then, some sort of stimulus is introduced which causes the participant to experience a state of temporal compression. While in this state of temporal compression, we may either ask the participant to (a) verbally estimate the amount of time that passed during their baseline interval (in which case they will simply say “five seconds”), or (b) *reproduce* the baseline interval that they experienced while in a state of temporal compression. In the latter case, because they are in a state of temporal compression, five objective seconds pass faster than our participant realizes. In this instance, *overestimation* therefore indicates temporal compression effects and *underestimation* therefore indicates temporal dilation effects.

In short, whether or not temporal estimation is prospective or retrospective and whether or not the estimation itself is coming from a baseline or altered temporal state directly determines how experiment results are interpreted. This complex interplay is summarized in the following figure (Figure 2.14):
Figure 2.14: A breakdown of how different temporal estimation methodologies require different interpretations. The temporal estimation paradigm featured in this dissertation constitutes an instance of ‘Altered Judging Baseline’ estimation.

Having now firmly and clearly established where the temporal estimation paradigm featured in this dissertation stands methodologically relative to previous research dealing with temporal perception, this allows us to move forward with interpreting our results in future chapters without fear of any sort of misinterpretation. This will be revisited in experiments two and three (Chapters 4 and 5), which are dedicated to investigating the behavior of temporal simulation.

Given the myriad ways in which language and time interact, the temporal estimation paradigm constitutes an exciting and practical methodological advancement for the study of a heretofore-undocumented aspect of mental simulation. The following section summarizes the material covered in Chapter 2, and after that, a brief preview of Chapter 3 is provided.


2.4 Summary of Chapter 2

Chapter 2 has provided a broad overview of the central concepts to be featured in this dissertation. In it, three major topics were covered. First, background on simulation semantics was provided. In this area, it was shown that simulation semantics represents a veritable paradigm shift in the study of how the mind represents meaning, but this new paradigm is not without certain flaws. Specifically, multiple inconsistencies within the motor facilitation and interference paradigm (e.g., inconsistencies in result patterns across different studies, inconsistencies in how simulation is supposed to handle abstract meaning, etc.) have made it difficult to produce a definitive picture of how and when simulation works.

It was suggested here (and in Chapter 1, as well) that accounting for certain lexico-pragmatic factors (i.e., usage frequency, lexical class, temporal density, and others) in simulation semantics experiments could help to mitigate some of the inconsistencies discussed above. These lexico-pragmatic factors can be linked with simulation semantics experimental results via the theory of sub-event monitoring, which divides these factors up into three unique tiers which interact to produce results that can be quite complex, but better capture the highly variable behavior of mental simulations.

Finally, it was stated that this dissertation would feature two experimental methodologies: the motor facilitation and interference paradigm and the temporal estimation paradigm. The motor facilitation and interference paradigm is, of course, well established within simulation semantics and has well-documented shortcomings and advantages. Unlike the motor facilitation and interference paradigm, however, the temporal estimation paradigm constitutes a new methodological paradigm to be used in simulation semantics in order to measure the extent to which temporal experience is encoded in simulation-based representation.
2.5 Looking Ahead to Chapter 3

Experiment one (Chapter 3) constitutes the first study to be featured in this dissertation. In it, several Spanish predications are tested for their simulation effects within the motor facilitation and interference paradigm. Crucially, stimulus sentences were tracked for the number of mentions of their verb-object combinations, which was taken to constitute a measure of short-term frequency.

It will be demonstrated in experiment one that short-term frequency does indeed significantly predict simulation results, and, in fact, seems to differentially affect various predication types in motor simulation, including congruent and incongruent predications, as well as predications of several different grammatical pattern types.
Chapter 3
Short-Term Frequency, Grammatical Patterns, and Motor Simulation Effects in Spanish

3. Introduction

In Chapter 2, which comprised a conceptual overview of the material to be covered throughout this dissertation, the theory of sub-event monitoring was introduced. As stated earlier, sub-event monitoring serves to unite lexico-pragmatic factors with simulation-based processing, and it can be defined as follows:

(1) Sub-event monitoring: The degree to which the salient simulation-based features of a linguistic predication are inspected due to lexico-pragmatic factors.

The theory of sub-event monitoring suggests that simulations are not context-independent embodied representations of semantic content, but are instead highly variable patterns of neurological activation which change depending on a number of different lexical and pragmatic factors. These lexico-pragmatic factors determine how closely (or not closely) a simulation is inspected, which, in turn, modulates how simulation effects are processed. The experiment reported in this chapter asks whether short-term frequency is a modulator of sub-event monitoring and thereby of motor simulation effects. Secondarily, this experiment looks at a variety of different grammatical patterns (to be described in section 3.3, below) as modulators of simulation effects, as previous research has provided preliminary evidence that grammatical content may indeed modulate simulation effects, as well. In order to investigate this secondary factor, this study uses Spanish-language stimuli. The use of Spanish stimuli not only enables us to explore the role of verbal morphology in motor simulation, but also allows us to study certain tense/aspect/mood contrasts that are not found (or not nearly as widely utilized) in English.
particular, this study explores formal distinctions related to the subjunctive-indicative contrast and the perfective-imperfective contrast.

Framed in terms of sub-event monitoring, this experiment primarily addresses the interaction between mental simulation and usage frequency—in this case manifested as (a) the number of mentions of particular verb-object combinations over the course of the experiment, and (b) the number of mentions of other linguistic devices (i.e., grammatical patterns) over the course of the experiment. As per the predictions outlined in Chapters 1 and 2, we expect that novel (i.e., recently introduced) verb-object combinations and grammatical features will produce motor simulation effects of a different character from familiar (i.e., repeatedly mentioned) verb-object combinations and grammatical features due to differences in inspection during sub-event monitoring. These predictions will be described more thoroughly in Section 3.1 below.

It will ultimately be shown that (a) number of short-term mentions does indeed significantly predict and modulate motor simulation effects, and (b) that the configuration of these motor simulation effects reverses as frequency increases, potentially due to effects of differential rates of motor habituation. Because we expect that differences between congruent and incongruent conditions (as well as between various grammatical patterns) will manifest themselves here as interactions with frequency which, as hypothesized in Section 2.2.5, may change heavily in character (e.g., potentially reverse in configuration) depending upon frequency, referring to these relationships as simple interference or facilitation effects may not suffice. Given this, I abandon the terms facilitation and interference. Instead, I refer to these conditions as congruent and incongruent to better capture the fact that the simulation effects produced by a single predication (or predication type) may produce both facilitation and interference effects, depending upon various sub-event monitoring factors like, e.g., short-term frequency and grammatical pattern.
The following section (Section 3.1) outlines the major predictions set forth in this experiment. Following that, Section 3.2 provides background information about (a) the motor facilitation and interference paradigm, (b) the role of grammatical patterns in mental simulation, and (c) the role of short-term frequency in mental simulation. Section 3.3 gives information about the methods, items, and participants employed in this study. Section 3.4 reports the significant results seen in the data. Sections 3.5 and 3.6 offer a general discussion of the findings and some concluding thoughts. Finally, section 3.7 looks ahead to experiment two (Chapter 4) and provides a brief preview of that content.

3.1 Predictions

Having established the basics of sub-event monitoring above (as well as in Chapter 2), I offer the following predictions:

(2) **Prediction One (Short-term frequency effects):** Recently introduced content will be monitored more closely than repeatedly mentioned content due to activation effects (or lack thereof, in the latter case). Given this, we expect that the number of short-term mentions of both verb-object combinations and grammatical patterns will significantly modulate (i.e., significantly predict) motor simulation effects. Stated differently, it is predicted that there will be a significant main effect for short-term verb-object combination mentions, as well as a significant two-way interaction for verb-object combination mentions and grammatical patterns.

(3) **Prediction Two (Sentence congruity effects):** We expect to observe significant differences between congruent and incongruent conditions, but it should be noted that
these differences will not necessarily manifest themselves in different overall average response time values, but may instead surface via an interaction between short-term frequency, verb congruency, and verb direction. It may be the case, for example, that verbs appearing in the incongruent condition are generally less susceptible to short-term frequency effects than verbs appearing in the congruent condition (or vice versa).

(4) Prediction Three (Grammatical pattern effects): Finally, it is predicted that there will be significant differences between participants’ average response times in certain grammatical patterns. Given findings from previous inquiry in this domain (e.g., Bergen & Wheeler, 2010), I hypothesize that realis, imperfective tense/aspect/modality combinations (e.g., the present progressive indicative) will generate stronger motor simulation effects than irrealis, perfective tense/aspect/modality combinations.

These predictions will be revisited in Section 3.5, below, which constitutes a general discussion of the findings from this experiment.

3.2 Usage Frequency, Grammar, and Motor Simulation

In the following subsections, I offer basic background information on three different topics: (1) The motor facilitation and interference paradigm (Section 3.2.1), (2) the role of grammar in mental simulations (Section 3.2.2), and (3) the relationship between short-term frequency and mental simulation (Section 3.2.3).
3.2.1 The Motor Facilitation and Interference Paradigm

As discussed in Chapter 2, the motor facilitation and interference paradigm is founded on a landmark 2002 study by Glenberg and Kaschak, in which participants were presented with various English-language sentences (imperative sentences and sentences describing both abstract and concrete transfer). After reading these sentences, participants were asked to decide whether or not they “made sense” (ibid, p. 559) by pressing one of two response buttons. Crucially, these response buttons (Yes and No) were oriented such that they ran parallel with the sagittal plane of the human body, where one was placed closer to the body, and the other was placed farther from the body. It was anticipated that certain sentences would affect the motor movement involved in pressing response buttons (by speeding up or slowing down response times), as simulation-based representation was theorized to recruit diverse areas of the brain, including, most pertinently, the motor cortex. Glenberg and Kaschak did indeed find this effect, termed the Action-Sentence Compatibility Effect, and, in doing so, spawned a new research paradigm within cognitive science and linguistics.

Motor facilitation and interference studies since Glenberg and Kaschak have looked at a diverse array of phenomena, although these phenomena tend to be primarily lexical. That is to say, most motor facilitation studies look at some sort of semantic contrast between groups of words that are grounded in physical motion (e.g., words denoting manual motion toward the body, words denoting manual motion away from the body). Fewer studies, however, have looked at the role of grammatical content in modulating motor simulation effects (Madden & Zwaan, 2003; Kaup et al., 2007; Bergen & Wheeler, 2010; Anderson, Matlock, & Spivey, 2013; Parril, Bergen, & Lichtenstein, 2013), and even fewer have looked at grammatical patterns in languages other than English (see, e.g., de Vega, Moreno, & Castillo, 2013). Furthermore, none of the above studies have attempted to bring usage frequency into the fold, even though it is well
attested to be an important influence on linguistic representation and behavior (see, e.g., Bybee & Hopper, 2001; Bybee, 2007).

3.2.2 The Role of Grammar in Mental Simulations

As stated above, work within simulation semantics has predominantly focused on the meaning of predicate-argument complexes and how they can affect the character of mental simulations. Glenberg and Kaschak (2002), for example, looked at (among other things) concrete instances of transfer (e.g., *You delivered the pizza to Andy*) versus abstract instances of transfer (e.g., *You told Liz the story*). Richardson and colleagues (2003), similarly, looked at sentences depicting horizontal scenes (e.g., *The mechanic pulls the chain*) versus sentences depicting vertical scenes (e.g., *The balloon floats through the cloud*). Other studies have focused on lexico-semantic distinctions too. For example, Yaxley and Zwaan (2007) found that sentences describing a low-visibility scene (e.g., *Through the fogged goggles, the skier could hardly identify the moose*) better primed participants to identify a low-visibility picture. Correspondingly, high-visibility sentences primed identification of a high-visibility picture (e.g., *Through the clean goggles, the skier could easily identify the moose*). In short, the above studies and many others provide ample evidence that simulation effects are driven by the semantics of both main verbs and their accompanying nominal arguments.

How grammatical content fits into this picture, however, has yet to be fully clarified. In work outlining the *Embodied Construction Grammar* framework, Bergen and Chang (2005, 2013) suggest that grammar affects simulation in various ways. While major syntactic constituents like subject nouns and verb phrases are demonstrated to supply the simulation content (see, e.g., Bergen, Lindsay, Matlock, & Narayanan, 2007), grammar modulates how this content is simulated. For instance, grammatical aspect is demonstrated to affect simulation salience (see, e.g., Bergen & Wheeler, 2010, Anderson, Matlock, & Spivey, 2013; Parril, Bergen, &
Lichtenstein, 2013), as is the presence of negation (Kaup et al., 2007). In addition to modulating simulation effects, Bergen and Chang suggest that grammar may even sometimes contribute additional semantic content to be simulated. As a hypothetical example, witness or firsthand evidentiality marking might implicitly render the interlocutor a participant within a given simulation, while, e.g., nonwitness or nonfirsthand evidentiality marking would presumably not do so.

In light of the above, it appears that grammar plays a role in mental simulation, although the character and extent of that role are still somewhat unclear. Further, given the relative paucity of studies dedicated to looking at the role of grammar in mental simulation, it is reasonable to conclude that the effects of grammar on simulation-based content are under-investigated.

The theory of sub-event monitoring suggests that language users ‘inspect’ simulations to varying degrees due to a variety of lexico-pragmatic factors, and these differences in inspection can dramatically affect how language-induced mental simulation plays out. It may therefore be the case that certain grammatical patterns are monitored more closely simply by virtue of their syntactic complexity or frequency. Similarly, grammatical aspect may also have an effect on simulation, as it will be shown in experiment three (Chapter 5) that lexical aspect (see, e.g., Vendler, 1957) affects simulation, as well.

Although the effect of short-term frequency on mental simulation is the primary avenue of inquiry here, this experiment nevertheless also represents an attempt to elucidate how various grammatical patterns affect simulation-based representation. It will therefore also be shown later on that there does, in fact, appear to be a rough correspondence between pattern frequency and simulation effects. This finding will be discussed in Section 3.5, below.
3.2.3 Short-Term Frequency and Mental Simulation

The effects of frequency on language use can be observed over the space of a single exchange in discourse (see, e.g., Gundel, Hedberg, & Zacharski, 1993; Lambrecht, 1996; Cumming & Ono, 1997, Schegloff, 2007), and they can also be observed over centuries (see, e.g., Bybee, 2006, Section 5 for review). This experiment concerns itself with the former set of frequency effects. As discussed in Chapter 2, while long-term frequency may affect things like a word or phrase’s phonetic representation or even its meaning, short-term frequency is concerned with activation status, which represents how present in the consciousness of an interlocutor a given linguistic unit is. For our purposes here, short-term frequency is defined as the number of times a particular verb-object combination or grammatical pattern is mentioned over the course of a single experiment. Obviously, such a definition becomes inadequate immediately beyond the boundaries of this experiment, as real-world language behaves very differently from language seen in an experimental setting. Nevertheless, the repeated activation of a given referent or pattern is something that occurs ubiquitously in everyday discourse.

One important effect of short-term frequency concerns comprehension. Specifically, there is a well-documented effect known as ‘semantic satiation’ whereby after uttering or reading a word or phrase enough times in quick succession, a language user will come to see that expression as meaningless. Spivey (2007) suggests that this loss of meaning is due to neuronal fatigue. Specifically, clusters of neurons associated with a given word or phrase become fatigued due to sustained activation and thus become unable to communicate with other neurons, effectively deactivating the neurological channels used to connect a particular meaning with a particular phonological form. Although it should be acknowledged here that the effects of number of mentions upon mental simulation would not strictly constitute semantic satiation effects (for one, these is simply much more intervening material between exposures in this
instance), the ramifications that repeated activation over the short term have for mental simulation are nevertheless relatively straightforward. As stated earlier in Section 3.1, it is expected that recently mentioned items will be closely inspected in mental simulation, but as number of mentions increases, this inspection will decrease. In this case, it is presumably not because the representation of the thing being referred to has substantially changed since its first mention (in, say, the way that a representation of a word or concept may change over years), but rather, because sub-event monitoring does not take place, the content being simulated is less accessible or even inaccessible.

### 3.3 Methods

In order to answer questions about short-term frequency, this study employed a repeated-measures design. Participants were exposed to several critical trials \((n = 96)\), and although they never saw any verb form more than once during the study, the verb lemmas themselves were repeated, and some noun forms were repeated as well. Critical trials were built with six different direct objects and twelve different verbs\(^1\), listed in Table 3.1, below:

<table>
<thead>
<tr>
<th>Verb</th>
<th>English Gloss</th>
<th>Direction</th>
<th>Paired Object</th>
<th>Paired Object English Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>arrojar</td>
<td>to throw, give</td>
<td>away</td>
<td>la pelota</td>
<td>the ball</td>
</tr>
<tr>
<td>tomar</td>
<td>to take, catch</td>
<td>toward</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rechazar</td>
<td>to reject, refuse</td>
<td>away</td>
<td>la cerveza</td>
<td>the beer</td>
</tr>
<tr>
<td>beber</td>
<td>to drink, imbibe</td>
<td>toward</td>
<td></td>
<td></td>
</tr>
<tr>
<td>lanzar</td>
<td>to launch, throw</td>
<td>away</td>
<td>el balón</td>
<td>the ball</td>
</tr>
<tr>
<td>atrapar</td>
<td>to catch, trap</td>
<td>toward</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(^1\) The verbs used here were gathered via a small norming study where native Spanish speakers \((n = 10)\) were asked to decide whether various Spanish verbs indicated manual action toward the body, away from the body, neither, or both. Those verbs which had an agreement of 80% or higher were used to construct stimulus sentences.

\(^2\) With the exception of jalar, the verbs used in this study are cross-dialectally quite common in Spanish, and the speakers utilized in this study all responded that they were familiar with every verb presented in a post-test interview.
As a result of this setup, participants saw particular verb-object pairings multiple times (although the verb conjugations in each pairing did change every time). These pairings were tagged for the number of mentions in the experiment (each verb was mentioned a total of eight times). This was the primary measure of short-term frequency used in statistical analyses performed for this experiment. As stated earlier, it was expected that fine-grained simulation effects would diminish as number of mentions increased, as such action representations would become less and less necessary as they were repeatedly simulated over the course of the experiment.

Stimulus sentences were presented in several different grammatical forms, each corresponding to a different tense/aspect/modality combination (save for the periphrastic and morphological future constructions, which share features of tense, aspect and modality). The forms used are listed below (with example sentences below each item). It should be noted that the forms listed below range from low-inspection (i.e., high type frequency) patterns like the simple present and present progressive to high-inspection (i.e., low type frequency) patterns like the morphological future and the present subjunctive. This range of inspection degrees and type frequencies will be briefly revisited later in Section 3.5:

(5) Simple present

*Inés lanza el balón.*

‘Ines throws the ball.’ / ‘Ines is throwing the ball.’

(6) Present progressive

*Inés está lanzando el balón.*

‘Ines is throwing the ball.’
(7) Preterite  
Inés lanzó el balón.  
‘Ines threw the ball.’

(8) Imperfect  
Inés lanzaba el balón.  
‘Ines used to throw the ball’

(9) Morphological future  
Inés lanzará el balón.  
‘Ines will throw the ball’

(10) Periphrastic future  
Inés va a lanzar el balón.  
‘Ines is going to throw the ball.’

(11) Present progressive indicative (in a subordinate clause)  
Parece que Inés está lanzando el balón.  
‘It looks like Ines is throwing the ball.’

(12) Present progressive subjunctive (in a subordinate clause)  
Parece que Inés esté lanzando el balón.  
‘It looks like Ines might be throwing the ball.’

As discussed earlier in section 3.2.2, grammatical patterns are hypothesized to modulate simulation effects. For example, some previous simulation semantics studies have indicated that grammatical aspect affects simulation salience (Bergen & Wheeler, 2010; Anderson, Matlock, & Spivey, 2013; Parril, Bergen, & Lichtenstein, 2013). Given that Spanish offers many grammatical patterns and/or distributions thereof that English does not (e.g., the use of the subjunctive mood varies heavily between the two languages), this experiment presented a suitable opportunity for inductive inquiry into the simulation-based behavior of these different grammatical patterns.

Several independent variables were manipulated in the stimulus sentences employed in this study. As mentioned above, verb mention number (1-8) was utilized as a measure of short-term frequency. Other variables included: the implied direction of the main verb (toward or
away), the congruency of the verb/button press combination (congruent or incongruent), and, as discussed earlier, the grammatical pattern of the stimulus sentence (the eight patterns featured above). This, accordingly, yielded an 8 x 8 x 2 x 2 study, where independent variables consisted of grammatical pattern, mention number, direction, and expected effect.

Corresponding to the variables outlined above, several statistical items were of theoretical import to the study. Of primary importance was a main effect for short-term frequency (i.e., number of mentions), but higher-level interactions were viewed to be important, as well. In particular, a three-way interaction between button press congruence, verb mention number, and verb direction was considered important as frequency effects upon average response time could feasibly vary depending upon sentence congruity and verb direction (a similar finding was, for instance, reported in Bergen & Wheeler (2010)).

Another important area of statistical inquiry concerns two-and-three-way interaction terms involving particular grammatical patterns and expected simulation effects. If, say, imperfect sentences generate significantly different simulation effects from preterite sentences, this would be reflected in a two-way interaction between grammatical pattern and verb-button-press congruence. Much like the frequency-based three-way interaction mentioned above, a three-way interaction between grammatical pattern, congruency, and verb direction was also utilized in the statistical model, as it could be possible that significant differences between two different grammatical patterns may only appear with one verb direction in the study.

Finally, beyond those items mentioned above, several additional main effects and lower-order interaction items were also necessarily included in the statistical analysis. These included, for example, main effects for congruency, grammatical features, and verb direction. These items will be treated in detail in Section 3.6, below.
3.3.1 Participants

Thirty Spanish-English bilinguals from the greater Boulder area participated in the experiment for financial compensation. Four of these thirty participants were not included in the statistical analysis (one participant used two hands while giving responses, one was seen texting during the experiment, one failed to follow the instruction of going as quickly as possible, and one failed to inform the experimenter at the halfway point of the experiment, when experiment conditions were supposed to be changed). The remaining participants ranged in age from 18 to 45 years (mean = 27.71, SD = 8.57; two subjects chose not to answer this item). Twenty participants reported themselves as right-handed, five participants reported themselves as left-handed, and one participant reported himself/herself as ambidextrous. All participants had normal or corrected vision.

Participants were also asked about which dialect(s) of Spanish they spoke, and, as would be expected, responses varied. Responses included Castillian \((n = 5)\), Mexican \((n = 5)\), Chilean \(n = 2\), Peruvian \(n = 2\), Venezuelan \(n = 2\), Argentinian \(n = 1\), Colombian \(n = 1\), Salvadorian \(n = 1\), Puerto Rican \(n = 1\) and others (e.g., “South American,” a combination of dialects, or didn’t respond to the question) \((n = 6)\). In short, the data comprised an eclectic mix of varieties of Spanish, although, most notably, Caribbean varieties (Cuban, Puerto Rican, Dominican) were poorly represented.

3.3.2 Materials

96 critical stimuli were grouped with 64 non-critical stimuli to yield a total of 160 stimulus sentences. These stimulus sentences were broken up into four blocks of 40 trials each. Non-critical stimuli consisted of nonsensical sentences like Pablo nadaba el balón (Pablo used to swim the ball). The critical stimuli were divisible by four different variables:
(13) Grammatical pattern: simple present, present progressive, preterite, imperfect, morphological future, periphrastic future, present progressive subjunctive (in a subordinate clause), and present progressive indicative (in a subordinate clause)

(14) Sentence direction: toward or away

(15) Direction-of-button-press congruency: congruent or incongruent

(16) Number of verb-object combination mentions: 1-8

Each sentence had in its subject position a two-syllable proper name (e.g., *Carmen, Berto*, etc.), and each direct object appeared with a definite article. This yielded stimulus sentences that shared the same general syntactic frame (save for the sentences containing a subordinate clause) and with comparable syllable counts (again, save for the sentences containing a subordinate clause). As a result, participants saw numerous sentences following the general pattern of, “*Moises bebió la cerveza.*” (*Moises drank the beer*) or “*José está empujando la puerta*” (*José is pushing the door*).

For sentences with subordinate clauses, the matrix verb *parecer* (*to look, seem*) was used, as, according to Butt and Benjamin (2004), *parecer* may take either an indicative or subjunctive verb following it. Thus, this yielded contrasts like the following:

(17) *Parece que Inés está lanzando el balón.*
   ‘It looks like Inés is throwing the ball.’

(18) *Parece que Inés esté lanzando el balón.*
   ‘It looks like Inés might be throwing the ball.’

As might be expected, whether or not speakers accepted or rejected stimulus sentences like (17) and (18) above was a matter of personal opinion, dialect spoken, and context (or lack thereof) provided. This will be discussed further in the sections below.

3.3.3 Design and Procedure

Participants were tested in one session that lasted roughly thirty-five minutes from start to finish. Participants were seated at a desk with a laptop computer and response collection
apparatus and initially asked to fill out a brief survey of their linguistic and cognitive backgrounds. Following this, the main task commenced. Instructions, given in both English and Spanish, asked participants to decide whether or not the sentences they saw made sense and could depict a realistic scene (following Bergen and Wheeler (2010)). Following a brief block of practice trials \( n = 5 \), any questions from the participant about the experiment procedure were answered, and then participants were free to complete the first half of the experiment in self-paced fashion. Halfway through the experiment, participants were instructed to inform the experimenter, who then switched the orientation of the response collection apparatus so that the response buttons were in opposite positions. Following another brief block of practice trials \( n = 5 \), participants began the second half of the experiment, and finished it in self-paced fashion.

Each individual trial of the experiment began with a blank screen. The act of holding down the central yellow button on the response collection apparatus would cause a stimulus sentence to appear onscreen. As soon as the participant released the yellow button, the onscreen sentence would disappear (pressing the yellow button down again would not cause it to reappear), and participants would then press either the green button or the red button to record a judgment. After one of these response buttons had been pressed, the screen would briefly flash to let the participants know that the response had been received and that the participant had progressed to the next trial in the experiment.

The main dependent variable of this study was response time—specifically, the time between participants’ releasing the yellow button (the button used to display sentences) and pressing the red or green response buttons. The general narrative for simulation semantics studies in the motor facilitation and interference paradigm is that variations in these response times signal simulation-based motor interference or facilitation effects (depending on how effects are construed), but in the case of this experiment, interpretation of simple average response times
as predicted by main effects was replaced by a model that predicted average response times with interaction terms including short-term frequency.

    Crucially, as this experiment was driven by participants’ assessments of whether or not sentences made sense and could describe realistic scenes, only critical sentences in which participants pressed down the green button (and were expected to press down the green button) were accepted as part of the data pool. Thus, of a possible 2,726 responses, 2,110 were collected, and the number of responses collected for each participant ranged from 57 to 96 (mean = 81.11, SD = 10.59).

3.4 Results

    Responses under 200 ms as well as those over 1,000 ms were removed from the data. This resulted in the exclusion of 6.3% of responses from statistical analysis, leaving a total of 1,978 responses for analysis. A mixed-models analysis with crossed random effects for participant and item was performed. This analysis yielded one significant main effect and several significant interaction terms, each of which will be discussed in turn below.

    As this was an 8 x 8 x 2 x 2 study, there were four possible main effects: sentence congruence (two levels), verb direction (two levels), grammatical pattern (eight levels), and verb mention number (eight levels). Three of these four independent variables were not found to be significant. These include sentence congruency (Wald $\chi^2(1, n = 26) = 1.63, p = 0.2$), grammatical pattern type ($\chi^2(7, n = 26) = 4.13, p = 0.76$), and verb direction ($\chi^2(1, n = 26) = 0.01, p = 0.91$). It should be noted, however, that all of these items were found to achieve significance when part of higher-order interaction terms. The remaining independent variable, verb mention number, was found to significantly predict participant response time ($\chi^2(1, n = 26) = 12.53, p < 0.001$). This significant correlation is illustrated in Figure 3.1, below:
**Figure 3.1:** A significant correlation between verb mention number and participant response time. A general downward trend is apparent in the line of best fit.

In the above chart, a downward trend in participant response times can be seen, where over the course of the experiment, participants’ response times decrease, on average, by over fifty milliseconds. This is likely attributable to simple habituation effects: as participants progress through the experiment, they become used to the task and are thus, on average, able to make button presses more quickly. It should be noted that this general habituation effect is apparent in several of the higher-order interactions highlighted herein, as well.

Two two-way interactions were also found to be significant in the statistical analysis. First, there was observed to be a significant interaction between verb direction and sentence congruity ($\chi^2(1, n = 26) = 6.16, p < 0.05$). This interaction is illustrated in the following figure (3.2):
**Figure 3.2:** A significant two-way interaction between verb type and sentence congruity. Sentences with toward verbs show a greater distinction between congruent and incongruent conditions.

Above, we see that, overall, *toward* verbs elicit much stronger differences between button-press-congruent and button-press-incongruent sentences than away verbs do. Additionally, if we reframe the above figure in terms of button press location (rather than verb direction type), it becomes apparent that motor simulation effects are strongest in the away-button-press condition (see Figure 3.3, below):
As seen in Figure 3.3, within both toward and away button press locations, although average response times for congruent and incongruent verbs fall within each other’s standard error range, there is a clear distributional difference between toward and away location condition. Specifically, in the toward-location condition, congruent verbs have slightly higher average response times (516 ms) than incongruent verbs (512 ms), while in the away-location condition, congruent verbs have slightly lower average response times (513 ms) than incongruent verbs (523 ms). Taken at face value, this suggests that motor simulation effects may be more readily observed in one button location than in another, echoing findings of some previous motor facilitation and interference research (i.e., Bergen & Wheeler, 2010). However, it should be noted that the above picture of significant interaction effects between button press location and sentence congruity is not complete. Specifically, adding an additional factor for number of verb
mentions in the interaction term generates significant results, as well. This finding will be addressed later in this section.

A second significant two-way interaction was observed in the data, in this case occurring between grammatical features and number of verb mentions ($\chi^2(7, n = 26) = 19.91, p < 0.01$). Specifically, the average reaction time per grammatical features is significantly modulated by number of verb mentions. This is illustrated in the following figure (Figure 3.4):

**Figure 3.4:** A significant two-way interaction between grammatical pattern and verb mention.

In the above figure, a few striking features immediately present themselves. First, one again notices an overall downward trend in the data, which is reflective of the task habituation effects discussed above. Second is the patterning of the best-fit lines in the above figure. Specifically, a general pattern is observable wherein for those grammatical patterns which initially elicit the
highest average response times (e.g., the present progressive and the preterite), ultimately elicit the lowest average response times. In fact, if we track average response time by grammatical pattern by each specific mention of said grammatical pattern (rather than by verb mention), this pattern becomes even more apparent (Figure 3.5):

**Figure 3.5:** An even more exaggerated distribution of best-fit lines with differential frequency effects presents itself when we sort them by grammatical pattern mention instead of verb mention.

Above, those grammatical patterns which have relatively higher average response times at the beginning of the experiment will have the relatively lower average response times at the end of the experiment. Conversely, those grammatical patterns which have relatively lower average response times at the beginning of the experiment will have relatively higher average response times at the end of the experiment. This finding will be revisited in Section 3.6, below.
Finally, two statistically significant three-way interactions were observed in the data. First, grammatical pattern, verb congruency, and verb direction were shown to interact in significant fashion ($\chi^2(7, n = 26) = 17.07, p < 0.05$). This interaction is shown in the following chart (Figure 3.6):

**Figure 3.6:** A significant three-way interaction between grammatical pattern, verb congruency, and verb direction.

Discerning a clear pattern in Figure 3.6 proves a very difficult task. For example, some grammatical patterns (e.g., the imperfect) show little difference between all four possible verb direction and sentence congruity combinations. Others are similar save for one outlier (e.g., the morphological future and the preterite). Finally, others present chaotic and unpredictable
distributions (e.g., the present progressive). Nevertheless, when looking at individual contrasts between grammatical patterns, we see that there are a few significant differences which help to reduce some of the noise seen above. For instance, the morphological future significantly differs from several tense-aspect combinations (see Figure 3.7, below).

**Figure 3.7:** A chart showing significant differences between the morphological future and three different grammatical patterns: the present, present progressive, and preterite.

![Significant Differences Between Morphological Future and Present ($p < 0.05$), Present Progressive ($p < 0.01$), and Preterite ($p < 0.05$)](chart.png)

Above, we see that the morphological future significantly differs from the present ($t(25) = 2.44, p < 0.05$), present progressive ($t(25) = -3.31, p < 0.01$), and preterite ($t(25) = -2.68, p < 0.05$). As with the other findings above, these findings will be treated in greater detail later in Section 3.6.

The second three-way interaction observed showed a significant relationship between verb congruency, verb direction, and number of mentions ($\chi^2(1, n = 26) = 6.07, p < 0.05$). As illustrated in the following figure (Figure 3.8), we see a pattern wherein differential effects of short-term frequency can be observed, such that short-term frequency appears to significantly
differentially modulate different predication types depending upon their button press location and sentence-to-button-press congruity.

**Figure 3.8:** A statistically significant three-way interaction between verb congruency, verb direction, and number of mentions.

While the distribution of lines in the above figure appears rather complex at first glance, it can actually be explained by a pair of rules regarding button-location-based habituation effects and sentence-congruity-based habituation effects. Again, these patterns will be discussed in greater detail in Section 3.5, below.

### 3.5 General Discussion

Three different hypotheses were presented in Section 3.1 above. These hypotheses are reiterated immediately below:

(21) **Prediction One (Short-term frequency effects):** Recently introduced content will be monitored more closely than repeatedly mentioned content due to activation effects (or lack thereof, in the latter case). Given this, we expect that the number of short-term
mentions of both verb-object combinations and grammatical patterns will significantly modulate (i.e., significantly predict) motor simulation effects. Stated differently, it is predicted that there will be a significant main effect for short-term verb-object combination mentions, as well as a significant two-way interaction for verb-object combination mentions and grammatical patterns.

(20) **Prediction Two (Sentence congruity effects):** We expect to observe significant differences between congruent and incongruent conditions, but it should be noted that these differences will not necessarily manifest themselves in different overall average response time values, but may instead surface via an interaction between short-term frequency, verb congruency, and verb direction. It may be the case, for example, that verbs appearing in the incongruent condition are generally less susceptible to short-term frequency effects than verbs appearing in the congruent condition (or vice versa).

(21) **Prediction Three (Grammatical pattern effects):** Finally, it is predicted that there will be significant differences between participants’ average response times in certain grammatical patterns. Given findings from previous inquiry in this domain (e.g., Bergen & Wheeler, 2010), I hypothesize that realis, imperfective tense/aspect/modality combinations (e.g., the present progressive indicative) will generate stronger motor simulation effects than irrealis, perfective tense/aspect/modality combinations.

Turning to prediction one above, it is clear that there is a significant main effect for short-term frequency in the data, but what motivates this main effect is slightly less clear. There appears to be a general tendency for participants to register lower response times as the experiment
progresses. While one could, in theory, attribute this basic finding to diminishing sub-event monitoring on the part of the participants, a much more tenable explanation is that participants are simply experiencing task-habituation effects. That is, as participants repeat the same motion in the task, they become more and more efficient at doing so and thus do so more quickly. Calling our first prediction confirmed is therefore something of a dubious proposition.

In light of the above, it appears to be more prudent to look for interaction effects which include number of verb mentions as a factor. In this case, two (or more) categories of differing slopes would potentially suggest differences in sub-event monitoring going above and beyond simple habituation effects. For example, as discussed in Section 3.4 above, the two-way interaction between grammatical pattern and number of mentions was observed to significantly predict participant response times. This particular finding will be discussed in detail later in this section, but, for the time being, it does suggest that sub-event monitoring considerations affect short-term frequency (and vice-versa) in determining simulation effects.

Turning to the second prediction above, it was posited that differences between congruent and incongruent conditions may not readily manifest themselves as simple grand averages in participant response times, but may instead prove to be different when viewed through the lens of short-term frequency. In fact, it was observed that there was no significant main effect for sentence congruence in the data gathered. However, once sentence congruence was factored in with verb direction and short-term frequency, a significant result was indeed generated. This three-way interaction shown in Figure 3.8 is repeated below (Figure 3.9):
Figure 3.9: A statistically significant three-way interaction between verb congruency, verb direction, and number of mentions.

Three-Way Interaction between Congruency, Verb Direction, and Verb Mention ($p < 0.05$)

As stated earlier, while the above figure appears to have a complex configuration of best-fit lines, the behavior of these lines can actually be captured by two simple rules. They are provided in (22) and (23) below.

(22) **Rule One:** Participants habituate to away button presses more efficiently than they do to toward button presses.

(23) **Rule Two:** Participants habituate to congruent button presses more quickly than they do to incongruent button presses.

Returning to Figure 3.9 above, we can see that lines representing *toward* button presses (represented by red and orange lines) have, in general, shallower slopes than lines representing *away* button presses (represented by green and blue lines). Furthermore, looking at each of the aforementioned pairs of lines, it is readily apparent that in the congruent condition of either pair, a steeper slope is seen, representing more rapid habituation.

It should be noted that the above result constitutes an instance of differential short-term frequency effects upon different classes of linguistic items, although in this case it should be noted
that item classes vary by button press location and sentence congruity, which are not features that are implicated in the theory of sub-event monitoring. It is, additionally, very difficult to directly integrate this finding with the theory of sub-event monitoring, as the act of pressing a button should certainly have no effect on sub-event monitoring in simulation that precedes it. However, it may be the case that the act of pressing a button affects recognition of subsequent mentions of a given verb-object combination: when the action performed matches with the semantic content that preceded it, participants are more quick to recognize and integrate that action/content combination for use in future trials. Conversely, if participants perform an action which is incongruent with the content preceding it, it becomes more difficult for them to recognize and integrate that action. This conjecture falls in line with previous findings looking at the relationship between embodied action and language learning (see, e.g., Macedonia, Müller, & Friederici, 2010; Macedonia & Knösche, 2011), where researchers have found, for example, that associating certain lexemes with corresponding iconic gestures allows for more efficient learning than not doing so.

The general finding that incongruent sentences are less susceptible to short-term frequency effects is intriguing because it captures a new dimension along which to characterize simulation effects. Calling the effects described above ‘interference’ or ‘facilitation’ is obviously insufficient when both are distinct possibilities that depend on short-term frequency. It may therefore be the case that, with regard to short-term frequency, button-press incongruity is captured by relatively shallow slopes, while button press congruity is captured by relatively steep slopes.

In short, the data addressed above do seem to confirm our second hypothesis. Although no significant difference was seen in overall average response times for congruent and incongruent trials, once these congruent and incongruent trials were separated by their mention number and
verb direction, a significant and relatively straightforward pattern emerged. Crucially, this pattern demonstrates an effect whereby the frequency is seen to differentially affect critical stimuli depending upon their congruence and verb direction. As will be shown in subsequent chapters, this pattern of differential frequency effects is not unique to this experiment.

Finally, it can be observed that the data lend a small amount of support to prediction three above—namely, that tense, aspect, and modality features modulate simulation effects. There are two pieces of evidence that support this: (1) a significant three-way interaction between grammatical pattern, verb congruency, and verb direction, wherein different grammatical patterns are shown to significantly correlate with different distributions of verbs and button press locations, and (2) a significant two-way interaction between grammatical pattern and verb mention number. A chart depicting the three-way interaction given as Figure 3.6 above is repeated below as Figure 3.10:
Figure 3.10: A significant three-way interaction between grammatical pattern, verb congruency, and verb direction.

Three-Way Interaction between Grammatical Pattern, Verb Congruency, and Direction (p < 0.05)

Above, one notices several striking features. Readily apparent is the fact that responses in the ‘SubSubj’ (subordinate clause of the subjunctive mood) group are markedly higher, on average, than all of the other categories (546 ms as against 514 ms). However, it is doubtful that increased average response times in the SubSubj group are due to simulation effects. Rather, this category was most frequently rejected by experiment participants as not being grammatical or depicting a realistic scene. In fact, of a possible 312 participant responses to sentences utilizing this grammatical pattern, only 113 were collected, yielding a response rate of 36.2%, which was far lower than all other grammatical pattern groups (whose average response rate was 91.4%). This
lowered response rate may ultimately be attributable to the matrix clause verb that was used in these stimulus sentences. Specifically, according to Butt and Benjamin (2004), there exists a small class of matrix clause verbs in Spanish that license a subordinate clause verb appearing in either the indicative or subjunctive mood. One of these verbs listed by Butt and Benjamin was \textit{parecer (to appear, to seem)}. While it is true that some speakers accepted \textit{parecer} sentences with subordinate clause main verbs in both the indicative and subjunctive moods, the vast majority strongly favored \textit{parecer} sentences in the indicative mood, suggesting that Butt and Benjamin’s claim about the verb may be attributable to dialectal idiosyncrasies. Alternatively, it may simply be the case that in order for participants to have accepted \textit{parecer} along with a subordinate clause in the subjunctive mood, more context would simply have to be provided. In any case, however, the higher response times in this group are more likely a product of participant hesitation than a product of interference effects generated by mental simulation.

Other features within Figure 3.10, however, are not readily attributable to participant hesitation. Looking at the remaining data on the chart, one immediately notices the diversity of average response times both across and within different grammatical pattern groups. Within the imperfect, for example, there is little difference in average response times among the various combinations of verb congruency and verb direction (away verb/congruent = 511 ms, away verb/incongruent = 518 ms, toward verb/congruent = 522 ms, toward verb/incongruent = 521 ms). On the other hand, the present progressive shows wild variation between these same conditions (away verb/congruent = 523 ms, away verb/incongruent = 489 ms, toward verb/congruent = 510 ms, toward verb/incongruent = 530 ms) which appears to defy any sort of logical patterning. Beyond that, categories like the morphological future and the preterite can be demonstrated to display very strong simulation effects, but ostensibly only in one specific condition. In the case of the morphological future, the away verb/congruent condition yields, on
average, a response time of 476 ms, while the other three conditions cluster at 521, 521, and 513 ms. Similarly, in the preterite, the toward verb/incongruent condition has an average response time of 546 ms, while the other three conditions have response times of 494, 504, and 499 ms.

In previous work detailing the relationship between grammatical patterns and simulation effects (Bergen & Chang, 2005, 2013), grammatical patterns were hypothesized to affect simulation in a few different ways. Most obviously, grammatical constructions assist in designating the roles of different participants in a simulation. For instance, in the active-voice version of the ditransitive construction, we know that (absent any intervening material) the first postverbal sister will be some sort of object of transfer (be it concrete, abstract, metaphorical, etc.), and the second will be the recipient (actual or potential) of said object of transfer. Thus, roles in simulation in English (and other languages which determine grammatical relations via word order) are assigned based on syntactic information. In other words, in a potentially ambiguous sentence like *The man gave the boy a dog*, we unambiguously simulate the boy as the recipient of a new canine companion (and not a dog as the recipient of a new human friend in *the boy*) by virtue of the role designation function that grammatical patterns serve.

Beyond the non-controversial function of role designation, however, grammatical patterns are posited by Bergen and Chang to serve two additional functions in simulation. Firstly, they are alleged to evoke general schemas to be used in simulation-based representation. For example, in a sentence like *The dog barked its way down the alley*, by virtue of the grammatical pattern in which the roles dog, bark, and alley appear, we know that the dog is moving through physical space in an alley and thus simulate it accordingly, even though there is a clear mismatch between verb meaning and construction meaning, and the main verb of the sentence, bark, specifies nothing about movement through physical space. Secondly, grammatical patterns are also attested to modulate the perspective of simulations. To illustrate, an active-voice sentence
like *The dog bit the man* would more prominently highlight the dog in simulation, whereas its passive-voice counterpart, *The man was bitten by the dog* would more prominently highlight the man.

In short, grammatical patterns help to designate the various roles of participants in simulation, evoke general schemas for simulation, and also modulate which participants in a given simulation are foregrounded or backgrounded. However, any attempt to fit a framework like that described above to the data would quickly prove misguided. For instance, in their 2010 study, Bergen and Wheeler concluded that simulation-based differences between the present perfect and the present progressive could be attributed to grammatical-pattern-based modulation via aspectual marking, where the present progressive would select for the ‘ongoing’ portion of a simulation, and the present perfect would select for its end-state. However, looking at the two grammatical patterns above that are considered exemplars of imperfectivity (the imperfect and the present progressive) we see that the task of finding consistent patterns cutting across both groups’ average reaction times proves impossible. As highlighted earlier, average response times within the imperfect group tend to cluster close to one another, while in the present progressive, they vary to a large extent. Turning to the only grammatical pattern in the group that generally entails completion of an event, the preterite, we also see that it does not mirror the distribution of the present perfect in the work of Bergen and Wheeler. Given this, it cannot be concluded that Bergen and Wheeler’s results were replicated in any fashion in this experiment, although, admittedly, there are several factors (e.g., the use of repeated-measures design, the language in which the experiment was conducted, the grammatical factors tested) which outright eliminate the possibility of a true replication to begin with.

Taken at face value, the findings detailed above suggest that previous theories of the manner in which grammatical patterns modulate simulation are insufficient to account for potential variation in the data. However, it should be pointed out that this does not necessarily
preclude the possibility that grammatical patterns may still, in some fashion, modulate lexical simulation effects. In fact, this study has yielded evidence suggesting that grammatical pattern elements and lexical elements are treated fundamentally differently in simulation. To wit, the distribution of average user response times per each individual viewing of a given verb is quite different from the distribution of average user response times per each individual viewing of a given grammatical pattern. This is illustrated in the following two charts (Figure 3.11, Figure 3.12):

**Figure 3.11:** Average user response times for each verb that appeared in the study sorted by number of verb mentions. Note that all of these verbs’ averages decrease in (roughly) parallel fashion as the experiment progresses.
Figure 3.12: A reprint of Figure 3.5, above. Note in this case that there exists a general pattern where grammatical patterns which initially yield relatively high average response times ultimately yield relatively low average response times, and grammatical patterns which initially yield relatively low average response times ultimately yield relatively high average response times.

To recap a finding briefly mentioned in Section 3.5 above, there appears to be a general trend where those grammatical patterns that are initially observed to have a relatively high average response time will be ultimately observed to have a relatively low average response time. Conversely, those grammatical patterns that are initially observed to have a relatively low average response time will be ultimately observed to have a relatively high average response time. This is exemplified by the distribution of data in Figure 3.12.

This patterning is strikingly different from that in its antecedent, Figure 3.11, where the average response time per verb is tracked against that verb’s number of unique mentions in the experiment. In this case, no pattern similar to that featured in Figure 3.12 is visible. Instead, we see a clustering of downward-sloping averages wherein verbs whose initial average response times are higher relative to other verbs have final average response times that are still higher relative to
other verbs. Concomitantly, verbs whose initial average response times are lower relative to other verbs may have final average response times that are lower relative to other verbs. Other verbs may deviate from this pattern slightly, but, in general, the distribution of the data is quite different from that featured in Figure 3.12.

According to the theory of sub-event monitoring, more closely monitored grammatical patterns are less susceptible to frequency effects. In this instance, it may be that more familiar grammatical patterns are monitored less closely and therefore undergo more significant short-term frequency effects, while less familiar grammatical patterns are monitored more closely and do not undergo significant short-term frequency effects. In fact, looking at the array of different patterns, we do see that more familiar patterns (i.e., the present progressive, the preterite, and the simple present) generally do have steeper slopes (i.e., are more susceptible to frequency effects), while the less familiar patterns (the morphological future, the periphrastic future, the subjunctive) generally do have shallower slopes. This could, in theory, help to explain the enigmatic nature of the distribution featured in Figure 3.12, although such a speculative explanatory mechanism should, at this point, be interpreted with caution. Nevertheless, it will be shown in coming experiments in this dissertation that this pattern does repeatedly occur across different tasks and simulation modalities.

---

3 This familiarity was determined with a simple search in the Corpus del Español (Davies, 2002). The searchable grammatical patterns, from most to least frequent (followed by individual corpus counts), are: **Simple present** (1,191,971), **Imperfect** (443,182), **Preterite** (386,218), **Present Progressive** (107,727), **Morphological Future** (67,040) and **Periphrastic Future** (12,372). While this finding does not perfectly predict the distribution above, it does preliminarily suggest that less frequent grammatical patterns (i.e., the morphological and periphrastic future) are less susceptible to short-term frequency effects, while more frequent grammatical patterns (i.e., the present, the imperfect, the preterite, etc.) are more susceptible to short-term frequency effects.
3.6 Conclusions

The theory of sub-event monitoring suggests that usage frequency, in its various forms, will affect how mental simulation plays out. As discussed in Chapter 2, usage frequency is composed of different forms of frequency (i.e., type and token) which are measured over different time courses (e.g., the space of a single discourse, weeks, months, years, etc.). In this experiment, it was hypothesized that as short-term usage frequency increased (via number of mentions of particular verb-object combinations and grammatical patterns over the course of the experiment), sub-event monitoring would naturally decrease, leading to an altered configuration of motor simulation effects. Such patterns were indeed observed in the data, specifically in the three-way interaction between verb congruency, verb direction, and number of mentions. This three-way interaction demonstrated that the effect of number of mentions significantly varied depending upon congruency and verb direction, where effects were stronger for away-location and congruent button presses.

Tracking average participant response time by other features also revealed similar patterns of differential frequency effects. For example, when tracking participants’ average response times by grammatical pattern mentions, we see a pattern where those predications types which initially elicited the longest average response times relative to other categories ultimately elicited the shortest average response times relative to other categories. As suggested above, this may be construed as evidence that speakers approach familiar and unfamiliar linguistic items (in this case, grammatical patterns) differently in mental simulation. While at this juncture such claims are merely speculative, the remaining experiments in this dissertation will, in fact, directly address the issue of familiarity (as instantiated by long-term frequency (experiments two, three, and four) and canonical ‘expectedness’ (experiment three)). As stated earlier, the above pattern
of differential frequency effects on linguistic categories will be shown to be a recurring theme throughout this dissertation.

In general, it should be noted that while the distribution of the data in this study initially appeared rather opaque (i.e., it was difficult to ascertain any sort of clear pattern involving grammatical pattern average response times), accounting for short-term frequency rendered this data distribution somewhat more transparent. At a bare minimum, this experiment can be safely regarded as evidence that short-term frequency is playing a part in determining outcomes in simulation semantics studies. Given this, several general statements can be made about the motor facilitation and interference paradigm, as a whole.

Firstly, although it does not present a very feasible option because of the increased numbers of participants that would have to be recruited, it may make more sense for future studies to gather fewer responses per condition per participant and instead recruit more participants in order to counteract the loss of statistical power that results from there being fewer critical item responses gathered. Because the configuration of simulation-based results appears to change heavily as number of exposures within an experiment increases, this may have the effect of nullifying motor facilitation and interference paradigm results that may actually have been observed to be statistically significant were fewer responses gathered from each participant. Alternatively, it may simply be the case that simulation effects need to be conceptualized not as static response time averages, but as moving averages which are modulated by frequency. This position will be revisited and taken up in later chapters and, in particular, in Chapter 7.

Secondly, in instances where researchers may, for whatever reason, be drawing solely upon a subset of data for illustrative purposes, it is crucial that it be understood that said subset isn’t necessarily a reflection of all of the data. For instance, a researcher looking at a subset of participants’ responses from the middle of a study period might conclude that no simulation
effects exist, while, in reality, there may be highly salient simulation effects that may be occurring prior to (and/or after) the time at which the data in question was gathered.

Looking for a mechanistic explanation regarding why an increase in short-term frequency differentially affect grammatical patterns in a fashion like that seen in Figure 3.12 above, we may look to Spivey’s (2007) description of neural population fatigue as one potential driving factor:

One easy way to undo a relatively stable state in a dynamical neural system … is through fatigue. If a neural population code is continuously stimulated for a significant amount of time, one can naturally expect that the refractory periods of the individual neurons will accumulate in number and duration until it becomes quite difficult to substantially excite that population code for some time. This has been demonstrated in neural firing rates in monkeys, in human neuroimaging, and in neural network simulations. (Spivey, 2007, p. 18)

Applying the above to the theory of sub-event monitoring, it may be the case that when a given lexeme or grammatical pattern is repeatedly activated, the semantic content supplied by that lexeme or grammatical pattern is consequently inspected less closely, altering processing time and therefore heavily affecting simulation semantics results. Given the general narrative that simulation-based activation causes a brief period of interference effects followed by a period of facilitation effects, we would thus expect to see mildly increasing facilitation effects as short-term frequency itself increased. This is, of course, what was observed in this experiment, although future studies will be needed to verify the effect and to more fully determine its behavior and character.

3.7 Looking Ahead to Chapter 4

Although the results of this experiment are promising, they also raise several questions. Most pertinently, although short-term usage frequency appears to modulate simulation effects, the various effects of long-term frequency remain an open question. The effects of temporal semantics on simulation salience remain an open question, as well. Given that, (a) statistically
significant differences between grammatical patterns were found, and (b) these grammatical patterns varied in terms of the temporal information they conveyed (i.e., tense, aspect), it does appear that temporal and grammatical factors may significantly affect language-induced mental simulation.

The remaining three experiments featured herein are meant to build upon the above. Specifically, all three utilize at different ranges and types of long-term frequency in the experimental stimuli given to participants, and experiment four (a second motor study) also has a more limited number of short-term mentions for each condition. Further, experiment four specifically tests two prominent grammatical patterns which participate in the English dative ‘alternation,’ the ditransitive construction and the oblique goal construction.

The following chapter (Chapter 4) presents experiment two, which introduces temporal estimation and temporal density, and also looks at the effects of long-term frequency on temporal simulation.
Chapter 4: Exploring Temporal Simulation Effects

4. Introduction

This experiment constitutes a first foray into a new area of experimental research. Specifically, I posit that because temporal perception is a central and constant aspect of embodied experience, the characteristics of temporal experience must necessarily be encoded within simulation-based representation. In this study, several different means of linguistically expressing temporal contrast (e.g., grammatical aspect, lexical meaning, temporal adverbs, etc.) are investigated using the temporal estimation paradigm outlined in Chapter 2. Here I aim to illustrate (a) that the temporal estimation paradigm is indeed viable as an experimental methodology, (b) that different temporal linguistic features have discernible effects on temporal simulation, (c) building on the results of experiment one, that verb usage frequency alters the manner in which subjects run time-course simulations, much in the same way that it alters motoric simulation. Crucially, it will also be shown that (b) and (c), above, can be subsumed within the theory of sub-event monitoring outlined in Chapter 2.

To reiterate, sub-event monitoring is defined as the following:

(1) The degree to which the salient simulation-based features of a linguistic predication are inspected due to lexico-pragmatic factors.

In this experiment, the lexico-pragmatic factors posited to affect sub-event monitoring (and thus, simulation) are (1) temporal density (see Section 2.2.3) and (2) usage frequency (see Section 2.2.1). Both of these will factor heavily into the discussion in Section 4.6 below.

In the following section (Section 4.1), I provide background information concerning the relationship between temporal cognition and temporal density, and also discuss temporal dilation
and compression effects. In Section 4.2, the predictions of this experiment are set forth, and in Section 4.3, I provide a brief note about the relationship between the temporal estimation paradigm and effects of morphosyntactic complexity. Section 4.4 comprises an overview of the methods used in this experiment, and Section 4.5 reports the results gathered. In section 4.6, I provide a discussion which relates the findings of this experiment to its initial predictions. Finally, in sections 4.7 and 4.8, I conclude the experiment and discuss the next research questions that will be addressed in this dissertation.

4.1 Time Perception and Temporal Density

A central tenet of simulation semantics is that cognitive representations are not disembodied, built up from conceptual primitives, or abstract; rather, they are experiential gestalts (see, e.g., Langacker, 1987; Barsalou, 1999, 2003). For example, a person’s mental representation of the action of throwing would, according to simulation semantics, be an experiential amalgam of all of the instances of throwing undertaken or perceived by that person over the course of their lifetime. Thus, were we to activate this person’s representation of the action of throwing by presenting them with the linguistic item that refers to the action under consideration (i.e., the verb throw), we would not only expect to see the so-called ‘language centers’ of this person’s brain activate, but the actual areas implicated in throwing, as well (i.e., the motor cortex, the visual cortex, and others).

The sensation of the passage of time, like the act of throwing a ball or visually witnessing a scene unfold, should also be considered an aspect of embodied experience. Although temporal cognition is fundamentally different from, say, motor or visual cognition, our conscious experience is nevertheless firmly ensconced within it. However, unlike other sensory modalities like olfaction or gustation, which are associated with relatively distinct neural regions, the neural
mechanisms which underpin the human sense of time are multifarious. In other words, there is not one region of the brain dedicated to temporal processing. Rather, there are multiple regions of the brain which work in concert with one another to produce the sensation of the passage of time, and these regions change depending on the type of temporal processing that is happening (e.g., subsecond processing vs. processing of longer intervals) (see, e.g., Lewis & Miall, 2003; Mauk & Buonomano, 2004; Buhusi & Meck, 2005).

Turning from the neurological level to the experiential level, there is much experimental evidence that has converged to indicate that experienced duration changes heavily depending upon the amount of information processed in a given situation. For example, participants are observed to overestimate the temporal length of a previously presented stimulus if that stimulus is novel (Tse, Intriligator, Rivest, & Cavanaugh, 2004; Pariyadath & Eagleman, 2004), dynamic (Kanai, Paflen, Hogendoorn, & Verstraten, 2006), complex (Schiffman & Bobko, 1974), or filled with salient events (Poynter, 1989) (for a general review including the above references, see Eagleman & Pariyadath, 2009).

What I argue here is that because embodied experience also carries with it a temporal component which can vary substantially depending on the above factors, perceptually-grounded simulations must necessarily also have a temporal component that varies, as well. The construct used in this dissertation to capture this variability is called temporal density. As defined in Section 2.2.3, temporal density is the number of transitions within the representational space of a predication. Those predications which are more temporally dense are inspected more closely in sub-event monitoring, leading to observable differences in mental simulation effects. In general, we expect that predications composed of complex and/or novel information will cause relative overestimation (i.e., temporal compression), while predications composed of simple and/or routine information will cause relative underestimation (i.e., temporal dilation).
To illustrate, consider the difference between a highly novel, infrequent predication like the ninjas square-danced around the fountain and a routine, frequent predication, like the man was a father. Because the first sentence is (a) more temporally dynamic (it constitutes an eventive predication) and (b) more novel (i.e., its words are more infrequent), language users should, according to the theory of sub-event monitoring, inspect that predication relatively closely. Conversely, because the second sentence is (a) less temporally dynamic, and (b) less novel, language users should inspect that predication less closely. The ultimate result of this is that there should, in theory, be measurable differences in temporal simulation between these two predications.

4.1.1 On Subjective Time, Objective Time, Temporal Dilation, and Temporal Compression

Although it was stated earlier in Section 2.3.2, it should again be noted here that how temporal simulation effects are construed in this dissertation (i.e., that overestimation corresponds to temporal compression while underestimation corresponds to temporal dilation) represents, to some extent, a break from previous literature. For example, Kanai and colleagues (2006), using a similar methodology to the one used here, concluded that participants who overestimated in the task were experiencing temporal dilation effects rather than temporal compression effects. A similar conclusion was reached by Eagleman and Pariyadath (2009). As mentioned previously, the discrepancy is likely arising here because of some confusion between the concepts of subjective time, objective time, temporal dilation, and temporal compression.

Consider a hypothetical situation in which a participant is presented with a temporal interval, shown some sort of stimulus, and then asked to reproduce the initial temporal interval. Our participant necessarily experiences that tone as taking a certain amount of time, which we will specify as $n$ subjective seconds. Following our participant hearing this tone of $n$ subjective seconds, some sort of stimulus is then introduced, possibly changing the rate at which time passes.
for this participant. The crucial factor here, of course, is that our participant doesn’t consciously know that time could now passing at a different rate for them. When they are asked to reproduce the tone, the length of the tone in subjective time is identical, but in objective time, it may now be different.

To make this example more concrete, if we were to present a participant with a tone of one objective second, and we found that this participant, after some sort of intervening stimulus, reproduced the initial tone as being two objective seconds, this would indicate not that time was travelling more slowly than it was before for our participant, but that it was now traveling more quickly for them than they realized. As our participant hears the tone initially, it will subjectively feel to our participant like \( n \) seconds, and when they go to reproduce this tone, it will still subjectively feel like \( n \) seconds, even though the duration of objective time has effectively doubled. In other words, what our participant thought was one second of objective time in the tone reproduction task actually turned out to be two seconds of objective time. This constitutes a perfect example of temporal compression. One experiencing temporal compression may think that ten minutes have passed when, in reality, twenty minutes have actually gone by.

Subjective time is an experiential constant, while objective time is what appears to change around us. This is, of course, mind-bendingly counterintuitive. Nevertheless, if one subjective hour is experienced over an objective temporal interval that is greater than one hour, this is without a doubt an instance of temporal compression. If one subjective hour is experienced as shorter than one objective hour, this is temporal dilation (for more, please see Section 2.3.2). Having firmly established the manner in which our results will be interpreted, we can now turn to the major predictions of experiment two.
4.2 Predictions

Using the temporal estimation methodology discussed in Chapter 2, I examine temporal simulation effects using several critical contrasts in this experiment. These include the grammatical and semantic contrasts exemplified in (1-5):

1. Habitual events vs. Single events
   (The man drove to Michigan often vs. The man drove to Michigan once)

2. Temporary states vs. Permanent states
   (The chef was angry vs. The chef was French)

3. Single events vs. Stativized events
   (The carpenter built a chair vs. The carpenter was building a chair)

4. Longer temporal intervals vs. Shorter temporal intervals
   (The woman worked on the project for one year vs. The woman worked on the project for one hour)

5. High-speed events vs. Low-speed events
   (The car zipped around the bend vs. The car inched around the bend)

In Section 2.2.3, temporal density was defined as follows:

(6) The number of transitions in the representational space of a predication

Following this definition, I postulate that the first predication in each of the above pairs describes a state of affairs whose representation has greater temporal density than that of its counterpart predication. We first consider the contrast between habitual events and single events, as exemplified in (1). Habituality entails repetition (see, e.g., Langacker, 1997) and repetition implies the presence of transitions between successive events. Therefore, in the case of habitual predications versus non-habitual predications, we expect the former to be more temporally dense, and thus induce temporal compression effects (i.e., overestimation). We next consider the contrast between temporary states and permanent states (see, e.g., Olsen, 1997). As it is implicitly understood that a temporary state must have both a starting point and endpoint, it is hypothesized in this instance that these endpoints represent transitions which should increase the
temporal density of temporary state predications relative to that of permanent state predications. Turning to contrast (3), which involves the past progressive, according to the model of temporal density, single events are considered more temporally dense than stativized events because the progressive is a state selector, and transitions are thus necessarily eschewed in predications of this grammatical aspect (see, e.g., Bickel, 1997; Michaelis, 2011). In contrast (4), it is predicted that predications of a longer temporal interval should, in theory, have more transitions contained within that temporal interval, leading to a higher overall temporal density. This is similar to the prediction outlined in our first contrast between habitual and single events, above. Finally, the high-speed events depicted in contrast (5) should be more temporally dense than low-speed events, as they depict rapidly changing, dynamic scenes with a high number of transitions.

Given the above, I set forth the following predictions. In general, high-density predications are associated with a high degree of sub-event monitoring, and low-density predications are associated with a low degree of sub-event monitoring. Because high-monitoring predications are associated with increased processing (and given the rationale supplied above in Sections 4.1 and 4.1.1), we expect these predications to produce relative temporal compression effects, and, correspondingly, we expect low-monitoring predications to produce relative temporal dilation effects. Thus, in each of the above contrasting pairs, it is anticipated that the former member (i.e., more temporally dense member) will have a greater relative average deviation from its associated target time than the latter. This statement can be construed in several different ways. These predictions are captured by the following figure (Figure 4.1):
**Figure 4.1:** Theoretically allowable configurations between more temporally dense (left columns) and less temporally dense (right columns) deviations from target time.

In the above figure, we see in all three cases that the left column is greater than the right column, even though each chart represents a different possible configuration of average user responses.

In Chart A, we see that the left column shows a moderate temporal compression effect, while the right column shows a moderate temporal dilation effect. In Chart B, the left column shows a strong temporal compression effect, while the right column shows a very weak temporal compression effect. Finally, in Chart C, the left column shows a very weak temporal dilation effect, but the right column shows a very strong temporal dilation effect. In each case, however, the more temporally dense (i.e., left) column generates *relative* temporal compression effects, and the less temporally dense (i.e., right) generates *relative* temporal dilation effects.

As our second and final prediction, it is anticipated that within the fast-speed verbs vs. slow-speed verbs contrast, corpus frequency will be demonstrated to modulate simulation effects. As outlined in Chapter 2, the theory of sub-event monitoring suggests that predication types with closely monitored semantic representations will be less susceptible to frequency effects than predication types whose associated semantic schemas are less closely monitored. Accordingly, I postulate that as frequency increases, both fast-speed and slow-speed verbs will transition from compression to dilation effects, but high-speed verbs, because they are more temporally dense, will be less heavily affected by frequency than their slow-speed counterparts. This complex
relationship will be captured by a simple two-way interaction term between verb type (i.e., fast-speed vs. slow-speed) and verb frequency, with average deviation from target time as the dependent variable.

4.3 A Note on Temporal Simulation Effects and Morphosyntactic Complexity

As discussed at the outset of this chapter, the temporal estimation paradigm constitutes a new methodology within simulation semantics. At the heart of the temporal estimation paradigm is the task of listening to tones of varying temporal intervals and then reproducing them by pressing a button. For example, if a population of experiment participants is presented with a tone of one second in length, their mean reproduction of that tone should also be one second in length. Critically, however, the temporal estimation paradigm introduces some sort of stimulus to participants between tone presentation and tone reproduction. Thus, if our participants’ average tone reproduction time deviates significantly from the tone presentation time, we can conclude that something about that stimulus material intervening between tone presentation and tone reproduction caused that deviation.

What is being suggested here is that such deviations in average tone reproduction time as those mentioned above are attributable to mental simulation effects. In this case, it is hypothesized that in the act of linguistic comprehension, neural timekeeping resources are recruited in simulation-based processing, thus subtly affecting the rate at which language users perceive time to be passing. It is also hypothesized that these temporal simulation effects are short-lived, much like the effects of visual or motor simulation, and last between a few hundred and a few thousand milliseconds. This dovetails with certain cognitively motivated theories of temporal perception, which suggest that the conscious awareness of time is segmented into a series of consecutive of two-to-three second windows (see, e.g., Chafe, 1994; Pöppel, 1994; for
review, see Evans, 2003). As the temporal simulation effects that occur within these windows are not only short lived, but also part of everyday subjective experience, they are, by definition, not perceivable. The temporal estimation paradigm, however, aims to show that such perturbations in temporal experience do, in fact, occur.

As stated above, deviations in average tone reproduction time are posited to be indicative of temporal simulation effects. A potential objection to this method, however, is that such deviations may not be the result of recruitment of neural timekeeping resources, but instead simply emerge due to stimulus sentences differing in morphosyntactic or semantic processing difficulty (see, e.g., Coll-Florit & Gennari, 2011). For example, a more difficult-to-parse sentence (e.g., *The state in which Paul was was embarrassment*) could cause participants to misestimate in their tone-reproduction responses. We do, in fact, find that this is, to some extent, true (see Section 4.5, below). However, multiple regression analysis allows us to factor in measures of parsing difficulty in our analysis, thus giving us a clearer picture of the extent to which deviations in average tone reproduction time are a product of simulation effects versus morphosyntactic complexity effects. Furthermore, contrasting pairs of sentences used in temporal estimation experiments generally differ by only one key feature. For instance, in an experiment looking at the effects of tense on temporal simulation, one might compare a sentence with the present-tense *hikes* with the past-tense *hiked* (e.g., *The boy scout hikes* vs. *The boy scout hiked*). In this case, the two stimulus sentences could differ by only one morpheme—the present form versus the past-tense form. It would be difficult, here, to attribute differences in average tone reproduction time to morphosyntactic processing difficulties, unless we were to posit that the preterite is inherently more difficult to process than the present tense.

Things become slightly more complicated when we are making comparisons between two different groups of stimulus sentences that differ in more than just the signification of a single,
crucial morpheme. For instance, in making an aspectual comparison between the simple present and the present progressive, the syntactic structure of the latter is somewhat more complicated, with there being an extra syntactic node, i.e., an auxiliary verb, situated (canonically) between the subject and the gerundial form. Although there are many ways to capture morphosyntactic complexity (for review, see, e.g., Wasow, 2002), syllable count is used here (primarily for its simplicity) (see, e.g., McDonald, Bock, & Kelly, 1993; Rosenbach, 2005) as the means of measuring the effect of morphosyntactic complexity on average tone reproduction time.

The following section reports on the experimental methods used in order to test the predictions set forth in Section 4.2.

**4.4 Methods**

In the following sub-sections, I outline the participants (Section 4.4.1), materials (Section 4.4.2), and design and procedure (Section 4.4.3) featured in this investigation. Following this, I turn to the results of the experiment in Section 4.5.

**4.4.1 Participants**

Sixty-six native-English-speaker undergraduate students enrolled at the University of Colorado Boulder participated in the experiment for extra credit. All sixty-six subjects successfully completed the experiment. These participants ranged in age from 18 to 24 years (mean age = 19.06). Fifty-six participants reported themselves as being right-handed, and ten participants reported themselves as being left-handed. All participants had normal or corrected vision.

Participants were also asked about a variety of other items. First, all sixty-six reported themselves as being native English speakers. Of those sixty-six native English speakers, five
reported having additional native fluency in one other language (languages reported were Spanish \((n = 2)\), Amharic \((n = 1)\), Gujarati \((n = 1)\), and Danish \((n = 1)\)). One speaker reported having native-level fluency in two languages other than English (Gujarati, Hindi). The remaining fifty-nine participants reported themselves as being monolingual.

Finally, all participants were asked to report whether or not they had a number of different conditions that might impact the experiment in some way. These included Dyslexia and Specific Language Impairment, as these conditions are implicated in reading task times (see, e.g., Botting, Simkin & Conti-Ramsden, 2006; Singleton, 2005) and Attention Deficit Hyperactivity Disorder, Schizophrenia, and Parkinson’s disease (as they are all implicated in temporal perception) (see, e.g., Gooch, Snowling, & Hulme, 2011; Davalos, Kisley, & Ross, 2002; Merchant, Luciana, Hooper, Majestic, & Tuite, 2008). While no participants reported themselves as having Parkinson’s disease, all of the other conditions were present in some form in the subject pool and were thus accounted for in the statistical analysis (see below) (ADHD \(n = 9\), Dyslexia \(n = 3\), SLI \(n = 1\), Schizophrenia \(n = 1\)).

4.4.2 Materials

Seventy critical stimulus sentences were grouped with seventy non-critical stimulus sentences for a total of 140 stimulus sentences used in the experiment. Stimuli were categorized along several variables: (1) linguistic features (habitual event vs. single event, unaltered event vs. stativized event, long interval vs. short interval, fast speed vs. slow speed, temporary state vs. permanent state, and filler), (2) length of accompanying tone (750 ms (filler only); 1,000 ms; 1,250 ms; 1,500 ms; 1,750 ms; 2,000 ms (filler only); 2,250 ms; 2,500 ms; 2,750 ms (filler only)), (3) number of syllables, and (4) trial block. This yielded seven critical trials per linguistic feature, and an average of 7.05 syllables per trial \((SD = 2.2)\).
Each pair of linguistic-feature contrasts was only seen with one accompanying time interval. While this made it difficult (but not impossible) to compare across different categories, it was determined that using multiple temporal intervals would yield a much greater variety of critical stimulus environments during the study itself, and thus improve the fidelity of the results, since presenting 70 of 140 trials as the same temporal interval would make the task very repetitive and likely cause habituation effects. Additional commentary on this decision is provided in Section 4.8 below.

Only critical trials contained a verb. Filler trials contained linguistic information that could, in theory, be simulated, except without a temporal component. These included phrases like fourteen potatoes, tennis instructor, and, a favorite of many participants, warlock inside of a dark tower. Trials were presented in a pseudo-randomized order such that two critical trials could be back-to-back (most were not), but no two consecutive trials were ever of the same temporal length. Lastly, all of the trials that could be marked for tense (thus, all critical trials) were presented in past tense. This was done in order to provide a degree of temporal consistency across critical trials.

As the comparison between high-speed and low-speed events was the only contrast where stimulus pairs contained different verbs, these verbs\(^4\) were tagged for their raw corpus frequencies using the Corpus of Contemporary American English (Davies, 2008-) for a basic analysis of frequency effects on temporal simulation. Their frequency scores are reported in Table 4.1 below:

\(^4\) Speed ratings of these verbs (i.e., fast speed or slow speed) were verified via a simple Mechanical Turk norming study where participants \(n = 21\) were asked to rate whether or not certain verbs (a) denoted fast movement, (b) denoted slow movement, or (c) neither of the above. Per-verb agreement between participants averaged 91.9\% (SD = 13.65\%).
Table 4.1: A list of all of the verbs used in the fast-speed/slow-speed condition with their groupings and frequencies.

<table>
<thead>
<tr>
<th>Verb</th>
<th>Group</th>
<th>Corpus Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Toddle</td>
<td>Slow Speed</td>
<td>192</td>
</tr>
<tr>
<td>Amble</td>
<td>Slow Speed</td>
<td>925</td>
</tr>
<tr>
<td>Trot</td>
<td>Slow Speed</td>
<td>2,417</td>
</tr>
<tr>
<td>Inch</td>
<td>Slow Speed</td>
<td>2,506</td>
</tr>
<tr>
<td>Glide</td>
<td>Slow Speed</td>
<td>3,200</td>
</tr>
<tr>
<td>Stroll</td>
<td>Slow Speed</td>
<td>4,299</td>
</tr>
<tr>
<td>Crawl</td>
<td>Slow Speed</td>
<td>8,509</td>
</tr>
<tr>
<td>Gallop</td>
<td>Fast Speed</td>
<td>1,198</td>
</tr>
<tr>
<td>Zoom</td>
<td>Fast Speed</td>
<td>2,264</td>
</tr>
<tr>
<td>Sprint</td>
<td>Fast Speed</td>
<td>2,327</td>
</tr>
<tr>
<td>Zip</td>
<td>Fast Speed</td>
<td>2,456</td>
</tr>
<tr>
<td>Dash</td>
<td>Fast Speed</td>
<td>3,027</td>
</tr>
<tr>
<td>Dart</td>
<td>Fast Speed</td>
<td>3,213</td>
</tr>
<tr>
<td>Charge</td>
<td>Fast Speed</td>
<td>30,290</td>
</tr>
</tbody>
</table>

Above, we see the list of verbs used in the fast-speed/slow-speed condition. A separate statistical analysis was performed on these items, but with *charge* omitted for its relatively high frequency. This will be discussed in greater detail in Section 4.6, below.

4.4.3 Design and Procedure

Participants were tested in one session that lasted roughly thirty minutes from start to finish. They were comfortably seated at a desk with a laptop computer and before undertaking the main task were asked to fill out a brief survey regarding their linguistic and cognitive backgrounds. After this, the main task was started. Instructions (given in English) stated that their goal was to reproduce tones as accurately as possible by pressing down the computer’s mouse button. A brief block of practice trials (n = 5) was presented, and, following that, participants were asked if they had any questions about the experimental procedure before beginning with the actual trials. After this, participants were free to complete the experiment in self-paced fashion. It was recommended to the participants that they take a brief break at the
halfway point of the experiment. Some participants chose to do this, while the majority did not. Following completion of the experiment, participants were debriefed by the experimenter as to the nature of their task and then were told they were free to go.

Each individual trial of the experiment would begin with the participant pressing down the space bar on the computer’s keyboard. This would initiate auditory stimulus (a 440 Hz sine wave tone) presentation. Following this, a sentence would appear on the computer monitor, and participants would read this sentence aloud. After this, participants would then hold down the mouse button for whatever length of time they estimated the auditory stimulus to have lasted. While doing this, the computer would generate the same tone as that which they had heard before in order to help participants with the task. After releasing the mouse button, participants would automatically receive a message saying, “Press the space bar to continue.” Doing so would initiate the next trial for the participants.

The dependent variable of this study was estimated time. Specifically, this was the length of time during which participants held down the mouse button, as mentioned above. This was also used to calculate two scores also used in the statistical analysis discussed below: raw deviation from target time and percent deviation from target time. Raw deviation from target time was used in statistical analyses because it was discovered that using percent deviation from target time distorted participant responses in higher target times. For example, a participant deviating by .1 s from a target time of one second will have a percent deviation of 10%, while a participant deviating by .1 s from a target time of three seconds will have a percent deviation of 3.3%. Using percent deviation in statistical analysis would, of course, be sensible if it were found that raw participant deviation proportionally increased as target time increased, but no such trends were found in this experiment (or in experiment three).
4.5 Results

Responses deviating from target times by more than one second were removed from the data. Thus, of 4,620 collected responses, 155 (or 3.4%) were removed. As in the statistical analysis that was performed in experiment one (Chapter 3), a mixed-models analysis with crossed random effects for participant and item was performed. The dependent variable used in the analysis was deviation from target time. This was calculated by taking participants’ individual response times and subtracting target response times from these numbers. Thus, overestimated responses are positive values, and underestimated responses are negative values. In the charts that follow below, adjusted deviation from target time is presented whenever possible. Adjusted deviation from target time takes into account effects generated by factors other than the grammatical features of a given condition (i.e., average syllable length of stimulus sentence and target tone length), and presents grammatical categories with those effects removed. In other words, the average deviation numbers presented in the adjusted charts better reflect the extent to which temporal semantics are affecting participants’ temporal cognition.

Looking at the data as a whole, we find several different significant main effects. Firstly, and most importantly, grammatical features (outlined in Section 4.2, above), were shown to very significantly predict average deviation from target time (Wald $\chi^2(9, n = 66) = 185.51, p < 0.0001$). This relationship is outlined in Figure 4.2, below.
Figure 4.2: An adjusted measure of participants’ response times for each grammatical category investigated in this study. Below, two pairs of contrasts are observed to achieve statistical significance.

![Adjusted Deviation From Response Time by Grammatical Category (p < 0.0001)](chart)

Figure 4.2 depicts the average deviation from target time for each of the five different critically contrasting pairs of sentence types. These contrasting pairs are adjacent to one another, beginning with the habitual event versus single event pair, and finishing with the permanent state versus temporary state contrast. The chart makes it apparent that a few particular critical contrasts, however, are unlikely to be significant. These include the preterite event vs. stativized preterite event contrast, the long interval vs. short interval contrast, and the permanent state vs. temporary state contrast.

However, two of these five pairs of contrasts were observed to significantly differ. The first, the habitual event versus single event contrast (Wald Z(64) = -1.85, p < 0.05), shows a weak compression effect for habitual events, as well as a strong dilation effect for single events. The second significant contrasting pair, fast speed verbs versus slow speed verbs, shows a significant
difference between each item ($Z(64) = 2.21, p < 0.05$), where fast-speed verbs generate a strong temporal compression effect, while slow-speed verbs generate a strong temporal dilation effect. A more detailed analysis of these findings will be provided in section 4.6 below.

One remaining significant main effect was discovered in the data as a whole. Specifically, there appears to be a significant main effect for ADHD status: those who self-reported as having ADHD are observed to slightly overestimate in their responses while those who did not were observed to underestimate their responses ($\chi^2(1, n = 66) = 4.88, p < 0.05$). These trends are illustrated in the following figure (Figure 4.3):

Figure 4.3: A significant main effect for ADHD status, where participants who self-reported as having ADHD are observed to underestimate significantly less than those who did not self-report as having ADHD.

None of the remaining independent variables (dyslexia status, age, bilingualism, syllables, order, number of syllables) were discovered to be significant. However, it should be briefly noted that dyslexia status approached significance ($\chi^2(1, n = 66) = 3.44, p = 0.06$). Also, the number of
syllables of a given sentence was associated with a general positive trend, but this trend was not observed to achieve significance ($\chi^2(1, n = 66) = 1.91, p = 0.16$).

As stated earlier, within the fast speed verbs versus slow speed verbs comparison, verbs were tagged for their frequencies using the Corpus of Contemporary American English. Here, several significant effects were observed. First, fast speed and slow speed verbs were again observed to significantly differ ($\chi^2(1, n = 66) = 5.25, p < 0.05$), reaffirming the results reported earlier. Second, a highly significant main effect for frequency was observed ($\chi^2(1, n = 66) = 20.43, p < 0.0001$). This main effect is summarized in the following figure (Figure 4.4):

**Figure 4.4:** *A highly significant main effect for frequency on the outcome variable of temporal estimation. Below, it is shown that low-frequency predications are associated with relative temporal compression effects (i.e., overestimation), and high-frequency predications are associated with relative temporal dilation effects (i.e., underestimation).*

Above, we see that as frequency increases, a general trend toward temporal dilation effects is observed. Disregarding the individual semantics of the verbs featured herein, this suggests that
more frequent verbs are less closely inspected during sub-event monitoring, leading to observed
temporal dilation effects (i.e., ‘boredom’ effects). This general negative trend in the data will be
discussed in greater detail in the following section.

Lastly, there was observed to be a highly significant two-way interaction between verb
type (i.e., fast speed versus slow speed) and frequency ($\chi^2(1, n = 66) = 18.11, p < 0.0001$). This
interaction is depicted in Figure 4.5, below:

Figure 4.5: A significant two-way interaction between verb speed and main verb usage frequency. Below, verbs
depicting fast-moving scenes are more sharply affected by usage frequency than verbs depicting slow-moving scenes.

Above we see depicted the significant two-way interaction between verb speed type and verb
corpus frequency where, as corpus frequency increases, a general trend toward temporal dilation
effects is present, but we also see what appears to be a pattern of convergence between the two
different groups of verbs. This will be discussed in greater detail below.
4.6 General Discussion

As discussed in Chapter 2, the theory of sub-event monitoring can be used to link simulation effects with lexico-pragmatic factors like usage frequency and temporal density. Indeed, in this experiment we see aspects of both usage frequency and temporal density affecting mental simulation. In the case of temporal density, we see that certain pairs of predication types that differ in their temporal density are significantly correlated with different average deviations from target response time, where more temporally dense predications are associated with temporal compression effects, and less dense predications are associated with temporal dilation effects. In the case of usage frequency, we see that as main verb usage frequency increases, relative temporal compression effects gradually become relative temporal dilation effects. Thus, we find confirmation of both of our initial predictions: (a) predications differing in their temporal density will generate significantly different temporal simulation effects, and (b) these temporal simulation effects will be modulated by frequency. In sum, temporal density and usage frequency are indeed crucial determinants in how language-induced mental simulation plays out within the temporal domain.

In addition, the experimental paradigm in which this study was conducted, the temporal estimation paradigm, appears to be a viable one. Although not all of the critical contrasts discussed above in Section 4.2 yielded significant results, two of these were found to be significant, suggesting that, with additional studies to verify its efficacy, temporal estimation may prove to be a fruitful methodology within the simulation semantics paradigm.

Briefly, as mentioned above, participants who self-reported as being diagnosed with ADHD significantly differed from their non-ADHD-diagnosed counterparts. In this case, the difference between subject groups was significant. In particular, while those without a diagnosis of ADHD tended to underestimate their temporal estimations by roughly 10 milliseconds, those
with a diagnosis of ADHD tended to *overestimate* their temporal estimations by a few milliseconds. This relative temporal compression effect may indicate that those with ADHD monitor sub-events in mental simulation more closely than those without ADHD. While this finding (and the theory of sub-event monitoring with which it is associated) is nothing more than preliminary, sub-event monitoring may prove a useful lens through which to investigate attention-based cognitive phenomena in the future.

Turning to the individual critical contrasts that varied in terms of their temporal density, we see that three of five pairs were not demonstrated to significantly differ. These included (1) temporary states vs. permanent states, (2) single preterite events vs. stativized preterite events, and (3) longer temporal intervals vs. shorter temporal intervals. These contrasts are depicted in Figure 4.6, below.

*Figure 4.6:* Critical contrasts investigated in experiment two that were not observed to significantly differ.

<table>
<thead>
<tr>
<th>Grammatical Features</th>
<th>Non-Significant Critical Contrasts (Adjusted Deviation)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Permanent State</td>
<td>Preterite Event</td>
</tr>
<tr>
<td>Temporary State</td>
<td>Stativized Preterite Event</td>
</tr>
<tr>
<td>Long Interval</td>
<td>Permanent State</td>
</tr>
<tr>
<td>Short Interval</td>
<td>Temporary State</td>
</tr>
</tbody>
</table>

![Graph showing non-significant critical contrasts](image-url)
Looking at the above, the question as to why the above three pairs of contrasts were not observed to significantly differ immediately presents itself. There are several possible reasons which may account for these observed results, each of which will be provided below.

Firstly, that no difference between preterite events and stativized preterite events was observed may ultimately be the product of referential ambiguity. In these stimulus sentences, semelfactive verbs were displayed in either the preterite or past progressive, yielding pairs of sentences like *the woman blinked* and *the woman was blinking*. Here it was originally hypothesized that a single preterite event would be more temporally dense than a stativized preterite event, as stativized predications are hypothesized to be less temporally dense than non-stativized predications. However, this prediction failed to account for the fact that predications like *the woman blinked* have multiple felicitous readings (i.e., *the woman blinked* can signify that she blinked once or blinked several times). Further, semelfactive verbs, when stativized via the progressive, are necessarily given an iterative reading (i.e., *the woman was blinking* cannot refer to a situation in which a woman was mid-blink, but rather, must refer to a situation in which she was blinking several times in succession). Given these issues (particularly with the preterite half of this contrast), it is perhaps not surprising that no significant result was observed.

In the case of long temporal intervals and short temporal intervals, in Section 2.2.3, it was suggested that changing the temporal scope of a predication (i.e., the size of a predication’s ‘representational space’) may also affect its temporal density. In this particular critical contrast, predications that were otherwise identical differed in the scope of a temporal adjunct they possessed (i.e., *for one hour* versus *for one year*). The temporal scope of a predication (i.e., the size of its representational space) may change, but the transition rate within that representational space could, in theory, stay constant, meaning that both predications, in this case, would still be of equal temporal density. Such a result doesn’t necessarily mean that temporal scope does not
affect temporal density, but in the case of when two otherwise identical predcations are being compared, it may not be a factor.

In the final critical contrast that was not observed to achieve significance, temporary states and permanent states were predicted to significantly differ because temporary states have starting points and endpoints (and should thus be more temporally dense), while permanent states do not (and should thus be less temporally dense). There are two possible explanations for why no significant result was observed in this contrast. First, it may be the case that even though temporary states have starting points and endpoints, these transitions are not featured within the representational space of temporary state predications. For example, after being told the chef was angry, one can feasibly ask, Is he still angry? or When did he become angry?, as this information is not implicitly represented within the representational space of this type of predication. Alternatively, it may be the case that the distinction between temporary and permanent states needs to be lexicalized for speakers to pay attention to this difference. This possibility is further explored in Section 4.8, below.

The two significant critical contrasts discovered in this experiment, habitual events versus single events and fast-speed verbs versus slow-speed verbs, are highlighted in the following figure (Figure 4.7):
**Figure 4.7:** The two pairs of critical contrasts from experiment two that were observed to significantly differ. Here, habitual events and fast-speed verbs were found to induce temporal compression effects, while single events and low-speed verbs were found to induce temporal dilation effects.

Above, we see two significantly differing critical contrasts, both of which conform to our theoretically allowable relationships between more temporally dense and less temporally dense material. In the above chart, the blue and green columns are composed of more temporally dense predications, while the red and purple columns are less temporally dense. In the case of the left pair, we see that the more dense of the two (habitual events) generates moderate temporal compression effects while the less dense of the two generates relatively strong temporal dilation effects. In the case of the right pair, we see that the more dense of the two (fast speed verbs) generates temporal compression effects while the less dense of the two generates temporal dilation effects.
It should be noted that in the case of the habitual event versus single event contrast above, simulation effects were generated by a temporal adjunct, and the adjunct in question was an adverbial phrase (i.e., *often* or *once*). This experiment is thus the first, to my knowledge, to demonstrate that simulation-based representation can be significantly modulated by adverbial semantics (as opposed to, say, verbal or nominal semantics).

In general, the compression and dilation effects seen in the habitual event versus single event contrast can likely be attributed to effects of temporal density on sub-event monitoring. Specifically, regarding habitual predication results, the stimulus sentences that elicited these results were eventive predications that were necessarily construed as repeated events spread out over a long time period by virtue of the added adverb *often*. I postulate that such repetition increases the temporal density of the predication, causing temporal compression effects. On the other hand, single events were unambiguously represented as such by virtue of the addition of the adverb *once*. In the case of events whose internal structures are repeated, because these events are simulated as taking place over and over again, this entails that there must necessarily be more transitions within the representational space of these predications. Conversely, in the case of events whose internal structures are unambiguously not repeated, it must necessarily be the case that these event representations contain fewer transitions than their repeated counterparts. As stated earlier, high-density predications are associated with relative temporal compression effects and low-density predications are associated with relative temporal dilation effects. Given this, the finding that habitual events are associated with relative temporal compression and single events are associated with relative temporal dilation fits perfectly with the predictions set forth in Section 4.2.

Within the fast-speed verb versus slow-speed verb critical contrast, it appears that fast-speed verbs generate temporal compression effects, while slow speed verbs generate temporal
dilation effects. Again, in this case, we would attribute the observed temporal effects to temporal density. Fast-speed predications depict dynamic scenes with high transition rates, and, accordingly, they should be inspected closely. Conversely, slow-speed predications depict static scenes with low transition rates, and thus they should not be inspected closely. As stated earlier, it is posited here that these differences in inspection are what drive the observed temporal compression and dilation effects. Again, the results pattern as predicted: high-density predications (i.e., high-speed verb predications) are associated with temporal compression effects, and low-density predication (i.e., low-speed verb predications) are associated with temporal dilation effects. However, as will be discussed below, these results should be interpreted with caution, as it will be demonstrated that verb speed significantly interacts with verb frequency, and these groups did not have identical frequency profiles.

We can observe two additional things about the fast-speed verb versus slow-speed verb critical contrast. Firstly, and most obviously, the contrasting grammatical element across stimulus sentences in different classes was the main verb. This yielded pairs of contrasting sentences like *the bird zoomed through the air* and *the bird glided through the air*. It is unsurprising that the lexical verb should be a major locus of contrast, as previous studies within the simulation semantics paradigm have found that simulation effects are strongly driven by main verbs (see, e.g., Richardson, Spivey, Barsalou, & McRae, 2003; Bergen, Lindsay, Matlock, & Narayanan, 2007).

Secondly, significant differences between fast speed and slow speed verbs, while broadly aspectual, are not in this instance attributable to lexical class differences, as traditionally construed. Using the heuristics outlined by Vendler (1957), as well as the premise that argument structure affects aspectual classification (see, e.g., Rappaport-Hovav & Levin, 1997), we can see a
relative homogeneity, in terms of Aktionsart class, across all verbs in the fast speed versus slow speed contrast. This is outlined in the following table (Table 4.2):

**Table 4.2:** A list of all of the verb phrases used in the fast-speed/slow-speed condition with their groups and grammatical aspects.

<table>
<thead>
<tr>
<th>Verb</th>
<th>Group</th>
<th>Grammatical Aspect</th>
</tr>
</thead>
<tbody>
<tr>
<td>toddle across the room</td>
<td>Slow Speed</td>
<td>accomplishment</td>
</tr>
<tr>
<td>amble across the field</td>
<td>Slow Speed</td>
<td>accomplishment</td>
</tr>
<tr>
<td>trot past the barn</td>
<td>Slow Speed</td>
<td>accomplishment</td>
</tr>
<tr>
<td>inch around the bend</td>
<td>Slow Speed</td>
<td>accomplishment</td>
</tr>
<tr>
<td>glide through the air</td>
<td>Slow Speed</td>
<td>activity</td>
</tr>
<tr>
<td>stroll down the street</td>
<td>Slow Speed</td>
<td>activity</td>
</tr>
<tr>
<td>crawl under the rock</td>
<td>Slow Speed</td>
<td>accomplishment</td>
</tr>
<tr>
<td>dash across the room</td>
<td>Fast Speed</td>
<td>accomplishment</td>
</tr>
<tr>
<td>charge across the field</td>
<td>Fast Speed</td>
<td>accomplishment</td>
</tr>
<tr>
<td>gallop past the barn</td>
<td>Fast Speed</td>
<td>accomplishment</td>
</tr>
<tr>
<td>zip around the bend</td>
<td>Fast Speed</td>
<td>accomplishment</td>
</tr>
<tr>
<td>zoom through the air</td>
<td>Fast Speed</td>
<td>activity</td>
</tr>
<tr>
<td>sprint down the street</td>
<td>Fast Speed</td>
<td>activity</td>
</tr>
<tr>
<td>dart under the rock</td>
<td>Fast Speed</td>
<td>accomplishment</td>
</tr>
</tbody>
</table>

Table 4.2 makes apparent that each predication’s grammatical aspect is evenly counterbalanced across verb speed conditions, thus eliminating the possibility that this may be driving the result seen. What remains, then, is the likely possibility that the differences observed between fast and slow speed verbs are the product of nuanced differences in verb meaning. For example, although a verb like *gallop* may pattern identically to a verb like *trot* in terms of Aktionsart heuristics, because the scenes of *galloping* and *trotting* are so different (the former is much more dynamic, the latter, much more predictable), this generates a reliable and significant difference between conditions. Although it will ultimately be shown that Aktionsart class does affect temporal simulation in the next experiment (Chapter 5), in this case, it appears not to be a factor in determining the distribution of the data.
Turning to frequency effects within the fast-speed verb versus slow-speed verb contrast, we see that significant results were generated here, as well. Once again, a few noteworthy features present themselves:

**Figure 4.8:** Figure 4.4, above, reprinted. Again, this figure demonstrates a highly significant main effect for frequency on the outcome variable of temporal estimation.

The above figure contains two salient features. First, there is a general negative trend in the data where less frequent verbs are more likely to generate relative temporal compression effects while more frequent verbs are observed to generate relative temporal dilation effects. It is again posited here that these effects are due to scalar differences in sub-event monitoring. However, unlike the pairs of critical contrasts discussed above which differed in terms of temporal density, what drives close inspection or non-close inspection in sub-event monitoring, in this case, is a given verb’s overall familiarity. Highly infrequent verbs are inspected closely, while highly frequent verbs are not. In this case, the result is the same: close inspection is associated with
relative temporal compression effects, while non-close inspection is associated with relative temporal dilation effects.

The second major feature presented in the above chart that merits mention is the clustering of its data points. As each verb featured in the above chart corresponds to a different corpus frequency, each ‘column’ of data points is representative of a single verb. Looking at these columns, however, it immediately becomes apparent that the general trend is for response times to be highly clustered, such that several verbs have standard deviations in the range of just a few milliseconds, even though differences between individual verb averages may be a second or more. This suggests that temporal simulation effects are both verb-specific and highly concentrated.

Looking at frequency effects across separate verb speed conditions, we are presented with a distribution of data that at once confirms and disproves our predictions (Figure 4.5, reprinted below as Figure 4.9):
Figure 4.9: A significant two-way interaction between verb speed and main verb usage frequency. Below, verbs depicting fast-moving scenes are more sharply affected by usage frequency than verbs depicting slow-moving scenes.

Above, we see that there is indeed a pattern of results where high-speed verbs and low-speed verbs are differentially modulated by frequency. However, going beyond this general pattern of differential frequency effects we see that it actually runs counter to our predictions. Specifically, the theory of sub-event monitoring states that high-monitoring predication types (in this case, fast-speed verbs) should be less susceptible to frequency effects than low-monitoring predication types (in this case, high-speed verbs). In other words, it was expected that in the above chart, fast-speed verbs would have a shallower slope, while slow-speed verbs would have a steeper slope, and this did not end up being the case.

There are number of explanations that we can posit regarding why the above distribution runs counter to predictions. However, absent additional data, these explanations are merely speculative. The first possible explanation for this distribution of data is that low-speed verbs and high-speed verbs have been incorrectly classified as low-density and high-density, respectively.
However, such post-hoc explanations are problematic for a number of reasons, and given the way in which temporal density has been defined (i.e., the number of transitions in the representational space of a predication), it is difficult to conceive of a way in which low-speed verbs could have more transitions than high-speed verbs.

Another explanation for the unexpected findings is that the range of frequencies sampled in this study is deficient. In fact, this is likely true on multiple levels. For one, the range of frequencies used here is rather small. The corpus counts represented in the above chart range from 192 to 8,509. To put that in perspective, the verb *speak* has a count of 135,511, and the verb *know* has a count of 1,035,794. Of course, the vast majority of verbs in the COCA corpus do not have such high counts, but, nevertheless, the sample used here does not fully capture the diversity of verbs available in the corpus. As a corollary to the above, not only is the range of verbs used here quite small, but all of these verbs are rather low in terms of their corpus counts. Considering again that verbs’ corpus counts in the COCA corpus range from just over zero to well over a million, a set of verbs ranging from 192 to 8,509 represents a rather low-frequency sample of verbs. Additionally, looking at the range of corpus counts for high and low speed verbs, one can see that they are not evenly counterbalanced, where the range of frequencies sampled for high-speed verbs is much smaller than the range of frequencies sampled for low-speed verbs. Essentially, it may be the case that were a larger, more balanced range of verbs used, the configuration of results seen could be completely different than that which was reported here. To return to a point made in Chapter 1, this is Hume’s problem of induction exemplified, and, admittedly, it could be the motivation for why these results do not conform to the theory of sub-event monitoring.

Turning to the final possibility for the above pattern not entirely conforming to our hypotheses, it may also be possible that low-speed verbs are inspected more closely because the
scenes they convey are less prototypical than those scenes depicted by fast-speed verbs (consider the difference between the team charged across the field and the team ambled across the field). This specific issue is, in fact, directly investigated in the following chapter (Chapter 5), yielding significant results, so it may indeed be the case that this is affecting the distribution of the data. Again, though, without future studies specifically dedicated to this issue, it is difficult to say with certainty what (if anything) is causing this pattern to not conform to our predictions.

4.7 Conclusions

Several important conclusions can be gathered from this study. First, in light of the results presented above, the temporal estimation paradigm does appear to be a viable experimental methodology within simulation semantics. From this, and given the multiple significant findings generated by this experiment, we can also tentatively conclude that there does appear to be a temporal component to language-induced mental simulation.

The claim that language-induced mental simulation has a temporal component should not be controversial. If simulation-based representation is built upon day-to-day experience, and day-to-day experience is always viewed through a shifting temporal lens, then it stands to reason that such shifting aspects of temporal experience would be encoded in simulation. As outlined in Section 4.1 above, it is believed that information density and the experience of the passage of time are inextricably linked to one another. For example, a situation which provides an utter lack of stimuli may trigger temporal dilation effects, while a situation which forces us to attend to a complex array of stimuli may trigger temporal compression effects. The most exciting finding from this experiment is arguably that linguistic predications that reflect these information characteristics produce measurable simulation effects that change the rate at which one time perceives to be passing, even if only temporarily and in imperceptible fashion.
In this experiment, we saw simulation effects generated by both temporal adverbs and main verbs. In the case of main verbs, we see that the effect distinguishing between fast speed verbs and slow speed verbs is not likely a product of Aktionsart class, but instead is a result of simple differences in meaning between the two different classes of verbs. It should be pointed out, however, that this finding certainly does not eliminate the possibility that Aktionsart class is a determining factor in temporal simulation effects, and this will, in fact, be investigated directly in experiment three.

Finally, this experiment provided evidence that two of the major factors posited to affect sub-event monitoring (and thus, mental simulation)—temporal density and usage frequency—indeed appear to affect simulation-based representation. Utterances varying in terms of their temporal density were demonstrated to significantly differ in terms of the simulation effects yielded. Furthermore, usage frequency (as instantiated by corpus counts) was observed to significantly predict temporal simulation effects. This finding falls in line with the findings of experiment one, which demonstrated that short-term usage frequency also significantly modulates simulation effects.

As predicted by the theory of sub-event monitoring, a pattern of differential frequency effects involving high-speed and low-speed verbs was observed to take place, although it should be noted that the specific details of this pattern ran counter to predictions. Nevertheless, it was once again demonstrated that failure to account for frequency can strongly effect simulation-based experimental results, as the relationship between the simulation results evoked by fast-speed and slow-speed verbs changed severely as frequency increased.
4.8 Looking Ahead to Chapter 5

Although the findings discussed above present a promising start, they also raise several critical questions. For one, there exists a conspicuous need to perform additional studies in order to verify the temporal estimation methodology as a viable one. While this methodology has thus far generated several noteworthy and significant results, future studies will be needed to ensure that such results are not, for example, statistical flukes.

One potential flaw in the design of experiment two is that its various critical contrasts (e.g., the temporary state versus permanent state contrast) were each only associated with one specific temporal interval (e.g., all temporary state versus permanent state sentences were associated with a tone length of 2.5 seconds, while all fast-speed versus slow-speed sentences were associated with a tone length of 1.75 seconds, and so forth). Thus, it is possible that something about the particular temporal interval with which each individual contrast pair is associated could be driving effects, i.e., either creating effects that aren’t actually there or diminishing effects that might be present otherwise. Experiment three will directly address this concern by looking at a number of different grammatical classes and features, with each displayed at multiple temporal intervals. If certain temporal intervals can be shown to generate contrastive effects, while other intervals do not, this will suggest that the results presented in this study are of diminished accuracy.

Another potential improvement involves corpus counts. Specifically, as discussed above, only a subset of the test items was tagged for corpus frequency (i.e., fast speed versus slow speed verbs only), and the range of frequencies looked at in the statistical analysis was relatively small and not balanced between conditions. Other verbs in the COCA corpus have substantially greater counts, and thus, a more representative sample is needed. Again, however, this issue will
be directly addressed in experiment three, where verbs of a much more diverse and balanced range of frequencies will be examined.

Finally, although the first two experiments have yielded several findings that align with the theory of sub-event monitoring as presented in Chapter 2, the theory does not capture all of lexico-pragmatic factors that may affect mental simulation. This is not only because we saw here one particular finding that certainly ran against the predictions of sub-event monitoring, but also because the theory itself simply has not been applied to a sufficient number of lexico-pragmatic factors and simulation modalities. The remaining two experiments serve, accordingly, as additional test beds for the theory of sub-event monitoring.
Chapter 5: 
Aktionsart Class, Usage Frequency, and Temporal Simulation Effects

5. Introduction

The work featured in this dissertation has thus far provided preliminary evidence that simulation-based representation interacts with a wide array of lexical and pragmatic factors, including, most pertinently, various forms of usage frequency. In Chapter 4, we explored the effects of frequency on temporal simulation, along with their implications for a version of simulation semantics in which understanding a depicted situation involves sub-event monitoring. The definition of sub-event monitoring provided in Section 2.2.5 is provided immediately below:

(1) The degree to which the salient simulation-based features of a linguistic predication are inspected due to lexico-pragmatic factors.

In Chapter 5, this exploration of frequency, temporal simulation, and sub-event monitoring is deepened. Specifically, Chapter 5 investigates whether or not different Aktionsart classes will generate differing temporal simulation effects, and also whether or not predications depicting highly expected and unexpected events will generate differing temporal simulation effects. Although the previous chapter tangentially dealt with issues of Aktionsart class (as a few of its pairs of critical contrasts differed in terms of their lexical aspect), it was not a direct area of inquiry. In Chapter 5, however, Aktionsart class will be prominently featured.

Before delving further into the main themes and predictions of Chapter 5, I will provide a brief review of the findings of Chapter 4. In Chapter 4, individual members of critical contrast sentence pairs were shown to significantly correlate with differing average deviations from target time, but the analysis of results raised some potential issues. Firstly, as the temporal estimation methodology has only been successfully implemented in one study thus far, it requires more
extensive validation. Secondly, the range of corpus frequency values looked at in Chapter 4 was not evenly distributed between groups or, for that matter, very representative of the different frequency values that one may find in a corpus: most of the verbs used in Chapter 4 were relatively infrequent. Furthermore, the only items tagged for their corpus frequency values were main verbs. In this experiment, other items (e.g., subject noun phrases) are tagged for their corpus frequency, as well. Finally, in Chapter 4, each pair of critical contrasts was associated with a particular temporal interval, raising the possibility that the effects observed were, to some extent, a product of the temporal intervals in which they appeared, as it could be the case that some temporal intervals are more conducive to generating significant temporal simulation effects than others are. In this Chapter, the temporal intervals used are counterbalanced across all conditions to insure that the effects observed are not a product of the temporal intervals with which they are associated.

To reiterate the findings of Chapter 4, two different pairs of critical contrasts were shown to significantly differ. These were habitual events versus single events (e.g., *The man drove to Michigan often* versus *The man drove to Michigan once*) and high-speed verbs versus low speed verbs (e.g., *The man sprinted down the street* versus *The man strolled down the street*). In both cases, the more temporally dense of each pair (habitual events and high-speed events) was shown to generate relative temporal compression effects, and the less temporally dense member of each pair was shown to generate relative temporal dilation effects. These findings are summarized in Figure 5.1, below:
Figure 5.1: Significant critically contrasting pairs from experiment two. Below, the more temporally dense member of each pair is shown to produce temporal compression effects, while the less temporally dense member is shown to produce temporal dilation effects.

Beyond the findings outlined above, there was also demonstrated to be a highly significant frequency effect for the fast-speed versus slow-speed verbs contrast wherein relative temporal compression effects were associated with low-frequency main verbs whereas relative temporal dilation effects were associated with high-frequency main verbs. Furthermore, fast-speed and slow-speed verbs were observed to be differentially affected by frequency, such that as frequency increased, the averages of these two groups began to converge and possibly even diverge.

However, as noted in experiment two, this pattern of differential frequency effects ran counter to the predictions set forth by the theory of sub-event monitoring, as the more temporally dense member of the pair was shown to be more susceptible to frequency effects instead of being less susceptible, as predicted.
Chapter 5 aims to resolve several of the ambiguities raised in experiment two while also aiming to reaffirm some of its results (and their posited explanations), as well. In particular, Chapter 5 looks at three major factors that should influence temporal simulation: (1) usage frequency, (2) expectedness, and (3) Aktionsart class. Each of these factors is discussed briefly in Section 5.2, below. Before that, however, I offer several predictions regarding usage frequency, expectedness, and Aktionsart class in Section 5.1.

5.1 Predictions

This experiment investigates several potential influences on sub-event monitoring and thereby on temporal simulation. Accordingly, several different predictions are set forth:

(17) **Prediction One:** In keeping with predictions made regarding temporal density and mental simulation, predications differing in terms of their lexical aspect (i.e., Aktionsart class) will be observed to significantly differ from one another regarding the temporal simulation effects they cause. Specifically, activity predications, which are the most temporally dense of the four Aktionsart classes analyzed, are predicted to evoke relative temporal compression effects, while state predications, which have the sparsest temporal representations, are expected to evoke relative temporal dilation effects. Achievements and semelfactives are both expected to fall between these two extremes.

(18) **Prediction Two:** In keeping with the predictions set forth in sub-event monitoring hypothesis, because highly expected events should be less closely inspected than highly unexpected events, it is anticipated that predications that depict highly expected events will cause relative temporal dilation effects, while predications that depict highly
unexpected events will cause relative temporal compression effects. Because closely-monitored predications are associated with a higher density of information processing, and non-closely-monitored predications are associated with a lower density of information processing, it is expected that expected and unexpected predications will be handled in much the same way as low-temporal-density and high-temporal-density predications, respectively.

(19) **Prediction Three:** Once again, frequency will have a significant modulating effect on simulation results: more frequent predications will generate relative temporal dilation effects, and less frequent predications will generate relative temporal compression effects. Furthermore, it is predicted that as frequency increases, the configuration of average response times for this experiment’s various conditions (in this case, Aktionsart classes) will change (i.e., diminish) and reverse due to different degrees of sub-event monitoring among classes. It should also be noted here that this prediction may potentially countervail prediction (1), above, as the discovery of a significant interaction effect between frequency and temporal simulation effects could markedly reduce the meaningfulness of any main effects in Aktionsart class that are discovered.\(^5\)

These predictions will be revisited later on in this chapter in Section 5.6.

### 5.2 Frequency, Time Perception, and Aktionsart Class

The following subsections present background information pertinent to the main themes of this experiment: usage frequency, time perception, and aktionsart class. As will be discussed

---

\(^5\) This possible outcome will be addressed in the discussion section (Section 5.5) later in this chapter.
later in the chapter, the theory of sub-event monitoring serves to unite these three relatively disparate items, such that usage frequency and Aktionsart class interact to produce distortions in participants’ time perception in the experimental task featured herein.

5.2.1 Usage Frequency

A high-level summary of the effects of frequency on language can be found in Chapter 2, but a few pertinent facts are presented here, as well. Because there are many ways in which frequency affects language, not all of these effects are pertinent here. Given this, the most relevant effects of frequency are discussed immediately below.

Of key importance to this study is the finding that usage frequency significantly modulates simulation effects. In particular, temporal simulation is affected. Results from experiment two (Chapter 4) demonstrate that processing low-frequency predications causes relative temporal compression effects (where participants are observed to overestimate relative to other conditions in the temporal estimation paradigm), and processing high frequency predications causes relative temporal dilation effects (where participants are observed to underestimate relative to other conditions in the temporal estimation paradigm).

In experiment one (Chapter 3), a significant effect of frequency was also reported. However, it should be noted that the type of usage frequency looked at in experiment one was fundamentally different from the usage frequency looked at in experiment two. Specifically, experiment one looked at short-term usage frequency (i.e., number of mentions over the course of the experiment), while experiment two looked at long-term (i.e., corpus) frequency. In both cases, frequency was observed to significantly differentially affect stimulus sentences appearing in different experimental conditions. The experiment featured in this chapter also looks at long-term frequency, although, as stated above, the long-term frequency used here is wider in range
and greater in depth. Specifically, a larger, more balanced sample of corpus frequencies was used for this experiment, and multiple types of frequency (i.e., main verb, subject noun phrase) were investigated.

5.2.2 Information Complexity, Expectedness, and Time Perception

In experiment two (Chapter 4), it was shown that more infrequent main verbs generated relative temporal compression effects, while more frequent main verbs generated relative temporal dilation effects. In other words, in this case, it was shown that a relatively novel (i.e., infrequent) stimulus would speed up time for participants, and a relatively mundane (i.e., frequent) stimulus would slow down time for participants.

This experiment has two different areas of investigation. The first looks at the effects of Aktionsart class and usage frequency on temporal simulation, and the second looks at the effects of canonical ‘expectedness’ on temporal simulation while controlling for word-level frequency effects. In the case of those trials in which canonical expectedness is investigated, subject-verb combinations are combined with typical and atypical direct objects, yielding contrasting pairs of stimulus sentences like the following:

(2) The messenger rode the bicycle.
(3) The messenger rode the dragon.

Because word-level frequency was controlled for as well as possible in this experiment (e.g., bicycle and dragon have relatively similar COCA Corpus frequencies of 5,214 and 4,811, respectively) any temporal simulation effects observed are more likely attributable to the effects of higher-level constituent frequency. In the case of the above sentences, for example, ride the bicycle is four times as frequent in the COCA Corpus as ride the dragon (although, admittedly, each have relatively low-frequency corpus counts of 4 and 1, respectively).
The theory of sub-event monitoring suggests that when speakers are confronted with novel information, they will inspect that information more closely in simulation, and when they are confronted with familiar information, they will inspect that information less closely in simulation. Unexpected predications, being more semantically atypical and of lower phrasal frequency, are therefore predicted to be inspected more closely in the temporal simulation task implemented in this experiment. In keeping with the findings of previous experiments highlighted above, it is anticipated that closely inspected (i.e., unexpected) predications will generate relative temporal compression (i.e., relative overestimation) effects and non-closely-inspected (i.e., expected) predications will generate relative temporal dilation (i.e., relative underestimation) effects. It should be noted here that this prediction falls squarely in line with previous research within temporal cognition which has shown that complex, unfamiliar, and salient stimuli cause participants to overestimate temporal intervals and simple, familiar, and routine stimuli cause participants to underestimate temporal intervals (Schiffman & Bobko, 1974; Avant, Lyman, & Antes, 1975; Thomas & Weaver, 1975).

5.2.3 Aktionsart Class

Although expectedness and frequency are important factors in this study, the role of Aktionsart class in temporal simulation is its main area of inquiry, as predications differing in their lexical aspect should, in theory, differ in terms of their temporal density. Among the significant and non-significant results seen in experiment two, it was shown that (a) some predications differing in their temporal density generated significantly different simulation effects, and (b) these effects were ostensibly strongest when they were generated by the main verb of a predication (rather than by an adverbial temporal adjunct). Accordingly, predications differing in their lexical aspect thus present themselves as a logical avenue for follow-up study. At stake
here are issues deeper than questions of validating a methodology or a particular type of simulation. In fact, this study is essentially investigating whether or not independently-motivated lexical classes exist insofar as their cognitive realization during mental simulation. Previous studies within simulation semantics (see, e.g., Richardson et al, 2003; Bergen et al, 2007; Bergen & Wheeler, 2010) have used ad hoc lexical classes (e.g., verbs denoting manual motion toward the body, verbs denoting manual motion away from the body), which, while semantically relevant to the task at hand, are not based upon distributional patterns of natural language. If different independently-motivated lexical classes were demonstrated to uniquely affect mental simulation, this would tell us that these classes are more than just a convenient means of describing groups of words, but instead reflect very real, distinctive cognitive categories that shape how we represent the world.

The modern concept of the Aktionsart class was posited by Zeno Vendler in his now-famous 1957 essay, ‘Verbs and Times’ (see also, Dowty, 1979). Vendler’s original taxonomy of Aktionsart classes included four classes: activities, accomplishments, achievements, and states. These are exemplified by the following predications:

(4) The athlete swam. (activity)
(5) The carpenter built a chair. (accomplishment)
(6) The train arrived at the station. (achievement)
(7) The farmer was Ukranian. (state)

Vendler astutely noted that, for example, stative verbs cannot appear felicitously as main verbs in questions beginning with the phrase At what moment, while achievement verbs could. Conversely, achievement verbs cannot appear felicitously as main verbs in questions beginning with the phrase For how long, whereas other Aktionsart classes can. These contrasts are illustrated below.

(8) At what moment did you arrive at the party? (achievement)
(9) #At what moment did you love the Beatles? (state)
(10) #For how long did you arrive at the party? (achievement)
(11) For how long did you love the Beatles? (state)
Vendler goes on to further delimit his Aktionsart classes by stating that accomplishments may felicitously appear with temporal durational phrases beginning with *in*, while activities may not.

(12) The carpenter built a chair *in* ten minutes.  (accomplishment)
(13) #The athlete swam *in* ten minutes.  (activity)

Since Vendler’s original work, other classes have been posited. Most notably, Comrie (1976, p. 42) put forth the class of verbs known as *semelfactives*. According to Comrie, semelfactives are punctual, like achievements, but differentiate themselves from achievements in certain situations. For instance, when used in the past progressive, a semelfactive verb like *blink* takes an iterative reading, while an achievement verb like *arrive* does not.

(14) The woman was blinking (repeatedly).  (semelfactive)
(15) The train was arriving at the station (#repeatedly).  (achievement)

It should be noted here that other scholars have since posited additional Aktionsart classes. For example, Van Valin and La Polla (1997) add active accomplishments and causative Aktionsart classes, as they note that activities can be given an accomplishment reading depending on the direct object taken by the verb, and that each of the above verb classes (save for semelfactives, as they are not in their classification system) has a causative variant that patterns differently from its non-causative variant. While such classes have their merits, they are simply beyond the scope of the current study, and will have to be investigated in future research.

Because the above Aktionsart classes differ inherently in terms of their temporal semantics, it is hypothesized here that they must also necessarily vary in terms of their temporal density, which was defined in Section 2.2.3 as the following:

(16) The number of transitions within the representational space of a predication

Regarding how one would go about solving the problem of representing temporal density, the work of Michaelis (2011) provides a useful framework for making distinctions between different
Aktionsart classes. According to Michaelis (as well as earlier work by Bickel (1997)), aspectual representations may be composed of two different components (either appearing individually, or in concert). These two components are τ and ϕ, where τ represents a transition or boundary, and ϕ represents a phase or state. According to the framework presented by Michaelis, a state would simply be represented as [ϕ], while an accomplishment would be represented as [τ ϕ τ ϕ] (for more, see Section 2.2.3).

Using the above framework, as well as the definition of temporal density provided in (16), we can categorize each of the five Aktionsart classes featured in study by their temporal density. Ordered from most dense to least dense, they are: (1) accomplishments ([τ ϕ τ ϕ]), (2) activities ([τ ϕ τ]), (3) achievements ([τ ϕ]), (4) semelfactives ([τ]), (5) states ([ϕ]). Given this, we expect more temporally dense predications to generate relative temporal compression effects (i.e., relative overestimation), while we expect non-temporally dense predications to generate relative temporal dilation effects (i.e., relative underestimation). As we shall come to see, however, the picture ultimately reveals itself to be a bit more complicated than that, thanks to the effects of usage frequency. We will return to this finding in Section 5.5, below.

To quickly reiterate the above, Aktionsart classes are believed to differ in terms of their temporal density, and it is hypothesized that predications differing in their temporal density will be simulated differently. This is because the theory of sub-event monitoring suggests that predications varying in terms of their familiarity and information density are handled differently in mental simulation. In keeping with findings from the previous two experiments, because usage frequency is also implicated in the theory of sub-event monitoring, where less frequent items are inspected more closely and more frequent items are inspected less closely, it is also hypothesized
that usage frequency will significantly modulate the simulation-based effects predicted by
differences in temporal density.

5.3 Methods

The following section composes a basic overview of the methods utilized in this study. In
it, I highlight the participants (Section 5.3.1), materials (Section 5.3.2), and design and procedure
(Section 5.3.3) featured in this experiment.

5.3.1 Participants

Fifty-two native English speakers over the age of eighteen years participated in the
experiment for either course credit or monetary compensation. All fifty-two of them successfully
completed the experiment. These participants ranged in age from 18 to 50 (mean = 22.71, SD =
6.74). Forty-five participants reported themselves as being right-handed, four participants
reported themselves as being left-handed, and three participants reported themselves as being
ambidextrous. All participants had normal or corrected vision.

As in the previous studies featured in Chapters 2 and 3, participants were asked about a
variety of other items. Firstly, all fifty-two participants reported themselves as being native
English speakers. Of those fifty-two native English speakers, only two reported having additional
native fluency in another language (languages reported were Spanish (n = 1) and Korean (n = 1)).
The remaining fifty participants reported themselves as being monolingual.

Finally, in keeping with the methodology of experiment two (Chapter 3), all participants
were asked to report on whether or not they had a number of different conditions that might
impact their results in some way. These included Dyslexia (n = 0), Specific Language
Impairment (n = 0), ADHD (n = 5), Schizophrenia (n = 0), and Parkinson’s Disease (n = 0). As
ADHD was observed in experiment two to significantly correlate with different response times, this item was included in the statistical analysis (discussed in Section 5.4, below).

5.3.2 Materials

Sixty-three critical stimulus sentences were grouped with sixty-five non-critical stimulus sentences for a total of 128 stimulus sentences used in the experiment. Critical stimuli were of two major types: (1) Aktionsart items ($n = 35$), and (2) expectedness items ($n = 28$). Five Aktionsart classes were examined in the study: (1) accomplishments ($n = 7$), (2) achievements ($n = 7$), activities ($n = 7$), semelfactives ($n = 7$), and states ($n = 7$). In the case of Aktionsart items, verbs were selected based upon their frequencies, such that the average frequency of each group of seven verbs was roughly equivalent. However, due to their having slightly different properties, accomplishment predications were not included in this frequency balancing. Specifically, this is because accomplishment predications in English are often a product of transitivity (for a very thorough treatment of this issue, see Rappaport Hovav and Levin (1997)). For example, the predication *The athlete swam* is interpreted as an activity, while *The athlete swam a mile* is viewed as an accomplishment as it has a specific goal and endpoint (cf. #*The athlete swam in twenty minutes* vs. *The athlete swam a mile in twenty minutes*). The accomplishment predications used in this study were thus created by taking an activity predication and adding a goal phrase to them. Given this, relatively complex verb phrases like *swim a mile* were necessarily more infrequent than relatively simple verb phrases like *swim*, making it impossible to felicitously include accomplishment predications in frequency balancing.

The COCA Corpus frequencies and standard deviations of each group are summarized in the following table (Table 5.1):
Table 5.1: A summary of the average COCA Corpus frequencies for each Aktionsart class used in this study. Also included are the standard deviations in frequency for each of these classes.

<table>
<thead>
<tr>
<th>Aktionsart Class</th>
<th>Average Frequency</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accomplishment</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>Achievement</td>
<td>26,961</td>
<td>20,205</td>
</tr>
<tr>
<td>Activity</td>
<td>28,680</td>
<td>28,895</td>
</tr>
<tr>
<td>Semelfactive</td>
<td>26,373</td>
<td>27,450</td>
</tr>
<tr>
<td>State</td>
<td>27,543</td>
<td>11,162</td>
</tr>
</tbody>
</table>

In the above table, the range of Aktionsart class frequencies employed in this study is provided. While the average frequency of each class hovers between 26,373 and 28,680, one notices quickly that the standard deviations listed above vary to a much larger extent (11,162-28,895). This is ultimately a product of the verbs that were available for study and their patterning in natural language. While other verbs from each class were available, the above sets of seven verbs constitute the most balanced possible groupings of Aktionsart classes using COCA corpus data.

In the case of the expectedness trials, because these stimuli weren’t revolving around differing lexical classes for verbs, all verbs were held constant across all conditions. They were wash, fold, cook, ride, shatter, whisk, and chew. Using the COCA Corpus, each verb was selected to appear with one direct object that was expected and (relatively) high frequency, one direct object that was unexpected but (relatively) high frequency, one direct object that was expected but (relatively) low frequency, and one direct object that was unexpected and (relatively) low frequency. For example, the verb *cook* appeared with direct objects pasta (7,388), sock (7,084), spaghetti (2,030), and stocking (2,399). Average frequencies and standard deviations for different types of direct objects are provided in Table 5.2, below:
Table 5.2: A summary of the average COCA corpus frequencies for the different categories of stimuli used in the ‘expectedness’ portion of this study. Standard deviations are provided, as well.

<table>
<thead>
<tr>
<th></th>
<th>Expected/ Frequent</th>
<th>Unexpected/ Frequent</th>
<th>Expected/ Infrequent</th>
<th>Unexpected/ Infrequent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average Frequency</td>
<td>11,783.14</td>
<td>11,253.57</td>
<td>2,136.43</td>
<td>2,999.42</td>
</tr>
<tr>
<td>SD</td>
<td>11,225.17</td>
<td>11,983.69</td>
<td>3,229.79</td>
<td>4,058.83</td>
</tr>
</tbody>
</table>

Subjects in all of these sentences types were held constant (within each verb), as well. For example, the verb *cook* appeared with *the chef* as its subject in each instance.

There were several other variables along which stimuli were categorized, as well. For instance, Aktionsart critical trials were also tracked for the temporal length of the interval tone with which they were associated (0.75 s, 1 s, 1.25 s, 1.5 s, 1.75 s, 2 s, and 2.25 s), the frequency of their subject (range = 310-67,051, average = 12,522.97, SD = 15,034.63), their number of syllables (range = 4-10, average = 6.8, SD = 1.89), their main verb, their presentation block, and order of presentation. In expectedness critical trials, stimulus sentences were also tagged for their subject frequency (range = 3,488-164,231, average = 36,610.43, SD = 54,653.94), syllables (range = 5-10, average = 7.07, SD = 1.41), and their presentation and block order.

Only critical trials contained a verb. Filler trials contained linguistic information that could, in theory, be simulated, except without a temporal component. These included noun phrases like *fourteen potatoes* and *tennis instructor*. Trials were presented in a pseudo-randomized order such that two critical trials could be back-to-back (most were not), but no two consecutive trials were ever of the same temporal length.

Lastly, all of the trials that could be marked for tense (thus, all critical trials) were presented in past tense. The purpose of this was to provide some degree of temporal consistency in the experiment.
5.3.3 Design and Procedure

Participants were tested in one session that lasted roughly twenty-five minutes from start to finish. They were comfortably seated at a desk with a laptop computer and before undertaking the main task were asked to fill out a brief survey requesting linguistic and cognitive background information. After this, the main task was started. Instructions were given (in English) which stated that their goal was to reproduce tones as accurately as possible by pressing down the computer’s mouse button. A brief block of practice trials \((n = 5)\) was presented, and, following that, participants were asked if they had any questions about the experimental procedure before beginning with the actual trials. After this, participants were free to complete the experiment in self-paced fashion. Participants were encouraged to take a brief break at the halfway point of the experiment. Some participants chose to do this, although the majority did not. Following completion of the experiment, participants were debriefed by the experimenter as to the nature of their task and then were told they were free to go.

Each individual trial of the experiment would begin with the participant pressing down the space bar on the computer’s keyboard. This would initiate auditory stimulus (a 440 Hz sine wave tone) presentation. Following this, a sentence would appear on the computer monitor, and participants would read this sentence aloud. After this, participants would then hold down the mouse button for however long or short they estimated the auditory stimulus to have lasted. While doing this, the computer would generate the same tone as that which they heard before in order to help participants with the task. After releasing the mouse button, participants would automatically receive a message saying, “Press the space bar to continue.” Doing so would initiate the next trial for the participants.

The dependent variable of this study was estimated time. Specifically, this was the length of time during which participants held down the mouse button, as mentioned above. This was
also used to calculate two scores also used in the statistical analysis discussed below: raw deviation from target time and percent deviation from target time. Raw deviation from target time was suitable for comparisons within critical contrasts, while percent deviation was suitable for comparisons both within and across critical contrasts, as those contrasts with longer estimation times were more likely to have a greater variance in terms of raw temporal deviation from the target time.

5.4 Results

Responses deviating from target times by more than one second were removed from the data. Thus, of 3,276 collected critical trial responses, 60 (or 1.8%) were removed. As with the statistical analyses done in previous chapters, a mixed-models analysis with crossed random effects for participant and item was performed. The dependent variable used in the analysis was raw deviation from target time, as classic ‘balanced’ outcome measures (e.g., ratio of recorded time to target time, or percent deviation from target time) were found to create a bias effect in the data where the range of values for lower target response times (e.g., 1 s, 1.25 s) was much greater than the range of outcome values for higher target response times (e.g., 2 s, 2.25 s). For example, if, regardless of temporal interval presented, participants tend to undershoot their estimations by one tenth of a second and we expressed this as a percent deviation, this underestimation would be -10% in a trial with a target time of 1 second, but a -5% deviation in a trial with a target time of 2 seconds. Given this, average raw deviation from target time was used in the statistical analysis.

Several significant main effects were found in the data. Beginning with expectedness trials, a significant main effect for expectedness was found (Wald $\chi^2(1, n = 52) = 7.08, p < 0.01$). This effect is summarized in the following figure (Figure 5.2):
**Figure 5.2:** A significant main effect for expectedness trials, where expected stimuli generate relative temporal dilation effects, and unexpected stimuli generate relative temporal compression effects.

In the above chart, predications with ‘expected’ direct objects are demonstrated to significantly differ from those with ‘unexpected’ direct objects, such that relative temporal dilation effects are observed in the former condition, while relative temporal compression effects are observed in the latter condition.

Statistical analysis also revealed a highly significant main effect observed for subject frequency \( \chi^2(1, n = 52) = 42.85, p < 0.0001 \) (Figure 5.3):
Figure 5.3: A significant main effect for subject frequency, where as subjects become more frequent, participants are more likely to overestimate in their responses.

A Significant Main Effect for Subject Frequency ($p < 0.0001$)

<table>
<thead>
<tr>
<th>Subject Frequency</th>
<th>Deviation from Target Time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>50000</td>
<td>0.0</td>
</tr>
<tr>
<td>100000</td>
<td>0.2</td>
</tr>
<tr>
<td>150000</td>
<td>0.4</td>
</tr>
<tr>
<td>200000</td>
<td>0.6</td>
</tr>
</tbody>
</table>

Although the above chart does indeed demonstrate a significant main effect for subject frequency (where more frequent subjects are associated with temporal compression effects), results should be interpreted with caution. In particular, of the seven different subject headwords used in expectedness critical trials, all but two had corpus frequencies under 25,000. The remaining two had disproportionately high frequencies of 44,547 and 164,231, and both were observed to be associated with temporal compression effects. Furthermore, as discussed above in Section 5.3.2, each subject was only associated with one verb (e.g., *The chef* was associated with the verb *cook*). Given this, it’s possible that this effect may simply be driven by main verbs, rather than their associated subjects. Thus, this evidence should not be taken as conclusive proof that subject frequency affects temporal simulation outcomes. While it may very preliminarily suggest that such a possibility could exist, this data, while highly statistically significant, is nevertheless too weak for one to draw reasonably safe conclusions.

Significant main effects were also observed within the Aktionsart trials. Before presenting these main effects, however, I remind the reader that accomplishment trials were omitted from
the analysis, as they differed too significantly from the other Aktionsart classes insofar as what they were tagged for (e.g., they could not be tagged for verb frequency). However, a separate analysis was done wherein activities and accomplishments were compared in terms of the overall simulation effects they generated. No significant difference was found in this analysis.

Turning back to our Aktionsart analysis, a significant main effect for Aktionsart class was observed ($\chi^2(3, n = 52) = 13.45, p < 0.01$). This effect is summarized in the following figure (Figure 5.4):

**Figure 5.4:** A significant main effect for aktionsart class, where activities appear to be associated with relative temporal dilation effects.

<table>
<thead>
<tr>
<th>Aktionsart Class</th>
<th>Deviation (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Activity</td>
<td>$-0.14$</td>
</tr>
<tr>
<td>Achievement</td>
<td>$-0.12$</td>
</tr>
<tr>
<td>Semelfactive</td>
<td>$-0.10$</td>
</tr>
<tr>
<td>State</td>
<td>$-0.08$</td>
</tr>
</tbody>
</table>

In Figure 5.4, four Aktionsart classes are depicted as causing temporal dilation effects. This figure, however, is somewhat misleading, because the averages of the different classes featured above have not been adjusted for certain factors (i.e., syllable and frequency effects). In light of this, a more thorough treatment of the above will be provided in Section 5.5, below.
Another significant main effect was discovered in the Aktionsart class data. This main effect was for verb frequency ($\chi^2(2, n = 52) = 14.24, p < 0.001$), and it is depicted in the following figure (Figure 5.5):

**Figure 5.5:** A significant main effect for main verb frequency. In this instance, less frequent verbs are associated with relative temporal compression effects, while more frequent verbs are associated with relative temporal dilation effects.

Above, main verb frequency is shown to be a highly significant predictor of temporal simulation effects. In particular, lower frequencies are observed to correlate with relative temporal compression effects, while higher frequencies are observed to correlate with relative temporal dilation effects.

Finally, the statistical analysis yielded a significant two-way interaction between Aktionsart class and verb frequency ($\chi^2(3, n = 52) = 14.36, p < 0.01$). This interaction is presented in the following figure (Figure 5.6):
**Figure 5.6:** A significant two-way interaction between Aktionsart class and main verb frequency. In this case, predications with state main verbs are most susceptible to frequency effects, while predications with activity main verbs are least susceptible to frequency effects.

The above interaction effect contains some very unusual and remarkable features, and will therefore be treated in detail in the following section.

As one final point, it should be noted that in both the Aktionsart and expectedness items, the other independent variables were determined to not significantly predict the data. These included ADHD status, number of syllables, block, and order of presentation.

### 5.5 General Discussion

This experiment has yielded several significant results, each of which is relevant to the major constructs and larger themes of this dissertation. Given the results highlighted above, it can be concluded that the temporal estimation methodology utilized here and in experiment two (Chapter 4) does appear to be one that can consistently produce results differentiating different types of linguistic predications. Not only have both temporal estimation experiments generated
significant results, but these results have also patterned in similar fashion. As a corollary to this, the results of this study also lend support to temporal simulation being a viable form of mental simulation. This theme and its antecedent will be discussed in detail in the following subsections.

5.5.1 Non-Significant Main Effects

It should be noted that while the major findings of this experiment fell in line with those of experiment two (Chapter 4), some of the less central findings did not. For instance, no significant effect for ADHD status was found in the data in this experiment. This may, however, simply be due to the fact that the number of participants with ADHD was somewhat lower in this experiment \((n = 5)\) as compared to the previous one \((n = 9)\). The overall number of participants, as well, was lower \((n = 52\) versus \(n = 66\) (previous experiment)). Because of this, the statistical power of such an analysis was likely diminished.

It should also be noted that in experiment two there was observed a general (albeit non-significant) trend whereby a larger number of syllables loosely correlated with an increase in relative temporal compression effects, and a smaller number of syllables loosely correlated with temporal dilation effects. However, in this experiment, no such trend was observed (Figure 5.7):
Figure 5.7: The relationship between number of syllables in a stimulus sentence and its average deviation from target time. A slight downward trend is visible, but this trend was not observed to achieve statistical significance.

In Figure 5.7, above, no positive relationship between number of syllables and deviation can be observed. In fact, in this case, we see a slightly negative relationship between the two. From this, we may draw the conclusion that this measure of morphosyntactic complexity does not appear to be a determining factor in temporal simulation effects.

5.5.2 Frequency

In this experiment, main-verb usage frequency was once again shown to have a significant modulatory effect on temporal simulation results. Specifically, low-frequency main verbs are associated with temporal compression effects, while high-frequency main verbs are associated with temporal dilation effects. This patterning is summarized in the following figure (Figure 5.8):
**Figure 5.8:** Average deviation per verb by that verb’s COCA Corpus frequency. In this case, the transition from temporal compression effects in low-frequency verbs to temporal dilation effects in high-frequency verbs appears to follow a fitted log-linear curve in relatively close fashion.

In Figure 5.8, a clear relationship between deviation and main verb frequency can be seen: as main verb frequency increases, deviation values increase toward the negative. Although the overall number of data points is relatively low, they appear to follow a log-linear distribution, rather than a simple linear distribution, suggesting that the extent to which usage frequency can affect mental simulation may approach a limit (in this case, an average deviation of roughly -0.25 seconds) as frequency increases. We would therefore expect that the difference between, say, a verb with a corpus frequency of, say 150,000 and another verb with a corpus frequency of 110,000 to be less than the difference between a verb with a corpus frequency of 50,000 and another verb with a corpus frequency of 10,000. Such a finding fits well with the theory of sub-event monitoring, in that there should be a limit to the amount of streamlining of inspection of sub-events that can take place in high-frequency predications while still retaining some form of simulatable meaning.
5.5.3 Expectedness Trials

Not only did experiment three generate significant results that, for the most part, have helped to verify the findings of experiment two, but it also generated several new findings. For instance, it was shown in the ‘expectedness’ trials that the canonical expectedness of a predication does indeed modulate its temporal simulation effects. In these trials, participants saw a subject NP + verb combination that was paired with expected, conventional object nouns and unexpected, unconventional object nouns (e.g, *The messenger rode the bicycle* versus *The messenger rode the dragon*). The results of these trials are summarized in Figure 5.2 (repeated as Figure 5.9, below):

**Figure 5.9:** A significant main effect for expectedness trials, where expected predications are associated with relative temporal dilation effects, and unexpected predications are associated with relative temporal compression effects.

Above, it can be seen that while both expected and unexpected trials cause temporal dilation effects, expected trials are observed to cause relative temporal dilation effects and expected trials are observed to cause relative temporal compression effects. These results pattern strikingly similarly to the verb frequency results discussed above. Namely, in both cases, less familiar (i.e.,
more heavily inspected) predications are associated with relative temporal compression effects, whereas more familiar (i.e., less heavily inspected) predications are associated with relative temporal dilation effects.

The above effect is noteworthy because it constitutes what is ostensibly a frequency effect, but not a frequency effect that takes place at the word level. As discussed above, object noun phrases were balanced for their frequency of occurrence in the COCA corpus such that for each instance of a verb appearing with an ‘expected’ direct object, an ‘unexpected’ direct object of similar frequency would also appear with said verb. The result of this is that the predications presented to participants differed in terms of the frequency of their higher-level constituents only. For example, whisk frequently co-occurs with egg, but much less frequently co-occurs with coffee. This suggests that there may be a complex interplay between different levels of frequency in determining mental simulation effects, such that two frequently occurring lower-level items could potentially have their word-level frequency effects nullified if they are presented as co-occurring and that co-occurrence is highly infrequent (e.g., think and eye are both highly frequent words, but predications about thinking eyes are highly infrequent). While further investigation of the relationship between word-level and higher-level frequency is a clear direction for future research, this topic is not treated elsewhere in this dissertation. However, in experiment four (Chapter 6), the complex relationship between two different forms of frequency, i.e., short-term frequency and constructional ‘preference,’ is investigated.

5.5.4 Aktionsart Class

Returning to new findings yielded by this experiment, an exciting new finding that has been observed here is that of significant differences between different Aktionsart classes in temporal simulation (Figure 5.4, reprinted below as Figure 5.10):
**Figure 5.10**: A significant main effect for aktionsart class, where activities appear to be associated with relative temporal dilation effects.

In Figure 5.10, we see average deviation times for the four Aktionsart classes used in the Aktionsart analysis (see section 5.4, above, for an explanation as to why accomplishment predications were not analyzed). It should be noted, however, that Figure 5.10 distorts the results, as each group of predications was associated with (1) a differing number of average syllables per stimulus sentence, and (2) a differing (albeit small) average main verb frequency for each group. After adjusting for those factors, we derive the following chart (Figure 5.11):
**Figure 5.11:** Average response times for each Aktionsart class group after being adjusted for their average number of syllables and average main verb frequency.

While the above chart does not starkly differ from Figure 5.10 above it, one can immediately note that stative predications are quite distinct from other Aktionsart classes. Correspondingly, activity predications are also more distinct.

Surprisingly, stative predications are observed, on average, to generate the weakest temporal dilation effects, while activity predications are observed to generate the strongest. It is quite important to note that, in general, this goes against what the theory of sub-event monitoring predicts should happen in temporal simulation. Namely, it is predicted that low-temporal-density predications should generate relative temporal dilation effects (much like high frequency predications), and high-temporal-density predications should generate relative temporal compression effects (much like low frequency predications). Crucially, however, the fact that the findings portrayed in Figure 5.11 run counter to predictions is not entirely problematic. Specifically, because there is a significant and complex two-way interaction
between Aktionsart class and main verb frequency, the findings portrayed in the above chart become much less powerful, such that were a different range of frequencies used in the experiment, it would, in fact, be possible to generate a configuration of temporal simulation effects that did not run counter to our predictions. This will be discussed immediately below.

The final significant finding generated by experiment three is a significant two-way interaction effect between Aktionsart class and usage frequency. This result is summarized in Figure 5.6, reprinted as Figure 5.12, below:

**Figure 5.12**: A significant two-way interaction between Aktionsart class and main verb frequency. In this case, predications with state main verbs are most susceptible to frequency effects, while predications with activity main verbs are least susceptible to frequency effects.

The above chart contains several striking features which merit mention and analysis. Most obviously, the above chart constitutes evidence that (a) there are indeed significant differences between Aktionsart classes insofar as the temporal simulation effects that they produce, and (b) frequency significantly modulates these temporal simulation effects, where low-frequency
predications generate relative temporal compression effects and high-frequency predications generate relative temporal dilation effects. However, exactly how these two facts manifest themselves in the data is something of a complicated issue that will be treated in detail in the following paragraphs.

Perhaps the most visible feature of Figure 5.12 is its extremely tight locus of convergence of average participant responses at a verb corpus frequency of roughly 30,000. Presumably, if one were to run a temporal simulation study wherein all of the verbs selected clustered around this corpus frequency, it is likely that no significant effects differentiating these verbs would be found. Similarly, as highlighted in previous chapters, if one were to run a temporal simulation study without taking frequency into account in any form, it is quite possible that no significant results would be generated, or, alternatively, that any significant results generated would not actually reflect the real distribution of the data. Using the above figure, consider, for example, a temporal simulation study that happened to utilize high-frequency states and low-frequency achievements. Its results would be striking and significant, but also, of course, highly misleading.

It is the above strongly modulatory frequency effect that, as mentioned previously, renders the simple response time averages for each aktionsart class all but moot. Again, had a lower range of frequencies been used, the collection of averages for the four Aktionsart classes featured here would have all skewed positive, but stayed in the same relative configuration as that seen in Figure 5.11 above. Conversely, if a higher range of frequencies had been used, those same Aktionsart class averages would have skewed even more negative, and would have likely appeared in the opposite configuration as that seen in Figure 5.11, and also would have confirmed our first prediction regarding temporal simulation effects, Aktionsart class, and average participant response time for each condition.
Turning back to the overall distribution of the data in Figure 5.12, we find that a pattern is again apparent wherein the simulation effects evoked by different linguistic classes (in this case, Aktionsart classes) are differentially modulated by frequency. This pattern of differential frequency effects has, by now, firmly established itself as a recurring pattern in the experiments presented in this dissertation. In experiment one, we saw two instances of short-term frequency differentially modulating motor simulation effects when sorting participant response times by factors like button press location, sentence congruity, and grammatical pattern. Similarly, in experiment two, we saw that temporal simulation effects generated by high-speed and low-speed verbs were differentially modulated by corpus frequency. Finally, in the case of this experiment, we see that the temporal simulation effects elicited by predications whose main verbs differ in their Aktionsart class are also differentially modulated by corpus frequency.

5.5.5 Reconciling these Findings with the Theory of Sub-Event Monitoring

According to the theory of sub-event monitoring, different lexico-pragmatic factors interact to produce complex simulation effects. As discussed in Chapter 2, there are three different ‘tiers’ of content involved in sub-event monitoring. The first tier consists of the lexical semantic content to be simulated. The second tier consists of frequency-based effects (i.e., short-term frequency, long-term frequency, etc.). Finally, the third tier consists of additional lexico-pragmatic factors (i.e., lexical class, temporal density, markedness, etc.) which interact with and modulate the effect of the second tier.

It was suggested in Chapter 2 that predication types which are associated with a high degree of inspection in sub-event monitoring will be less susceptible to frequency effects, and, correspondingly, that predication types which are associated with a low degree of sub-event monitoring will be more susceptible to frequency effects. And, indeed, that is what we see when
we look at the data. In Figure 5.12, for instance, we see that the steepest slope (i.e., the strongest frequency-modulated change in temporal simulation effects) belongs to stative predications, which, as outlined above, are lowest in temporal density of all four Aktionsart types investigated herein. Correspondingly, activity predications have the highest temporal density, and, accordingly, their change from relative temporal compression effects to relative temporal dilation effects is much more gradual as they move from low to high frequency.

Although the above constitutes a series of rules that effectively explain the patterning of the data in Figure 5.12, there nevertheless remains a question as to what cognitive factors or processes would motivate such a patterning. The answer proposed here involves processing speed (see, e.g., Forster & Chambers, 1973, Forster 1976; Balota & Chumbley, 1985; Ellis, 2002; Gennari & Poeppel, 2003; Ferretti, Kutas, & McRae, 2007; Coll-Florit & Gennari, 2011). In a 2011 study, Coll-Florit and Gennari found that durative predications like own a pool were processed significantly more slowly than punctual predications like fall into a pool. The authors posited that the effect observed could be due to two different possibilities: (1) durative predications represent scenes taking place over longer periods of time than the scenes depicted by punctual predications, and thus take longer to process, or (2) because durative states should, in theory, have a greater number of associated meanings and contingencies, they take longer to process.

It is difficult to characterize the effect observed by Coll-Florit and Gennari as a temporal simulation effect, as (a) the study itself revolved around a reading task (and not a task measuring temporal cognition), and (b) the authors provided no evidence that temporal perceptual resources were co-opted for simulation-based representation. Nevertheless, their findings do indeed suggest that predications differing in their temporal semantic content may be processed differently.
It is fairly difficult to characterize the data presented here as conflicting with Coll-Florit and Gennari’s findings, as the experimental procedures used in their study differ extensively from those used here. Of course, it is also difficult to say that their findings are entirely compatible with those presented here. Given this, I offer an alternative account of temporal processing which follows immediately below.

Recent work in cognitive science dealing with neural timekeeping mechanisms has shown that there is not one specific region of the brain that can be characterized as being specifically dedicated to timekeeping. Rather, timekeeping emerges as a product of the synchronicity between diverse neural regions (see, e.g., Lewis & Miall, 2003; Mauk & Buonomano, 2004; Buhusi & Meck, 2005). Given the neural ubiquity of both temporal processing and language processing, it should come as no surprise that there is some documented overlap between the two. In particular, there is overlap in neural regions like the cerebellum and the basal ganglia (Rao et al, 1997; Stevens, Kiehl, Pearlson, & Calhoun, 2007; Grosson et al, 2003; Booth et al, 2007). As such, it is posited here that linguistic processing co-opts and thus affects these neural timekeeping regions.

As discussed in Chapter 2, highly frequent words and phrases are processed differently than highly infrequent words and phrases. Frequent words and phrases, specifically, are processed more quickly. Their being referenced and used often has made them more rapidly and easily accessed. Less frequent words and phrases, on the other hand, are accessed slowly. They are often activated along with more frequent lexical neighbors, generating interference effects and slowing processing speed. It is posited here that because there exists a potential overlap between linguistic processing and temporal processing, the speed with which language is processed can affect the speed of one’s ‘internal clock,’ for lack of a better term. If language processing occurs rapidly, then neural timekeeping should also move at a rapid pace, thus
resulting in temporal dilation (i.e., the experience of time as passing more slowly than usual). Conversely, if language processing occurs slowly, then neural timekeeping resources would thus also move at a slow pace, resulting in temporal compression effects (i.e., the experience of time passing more quickly than usual). This model is compatible with the Droit-Volet and Meck (2007) model of neural timekeeping discussed in section 2.2.3, which states that relatively onerous processing tasks may also divert neural resources that are otherwise utilized in keeping track of time, while easier processing tasks would (in theory) do the opposite. In this case, we might build upon their model by suggesting that onerous processing tasks don’t divert away from timekeeping resources, but instead simply involve them: if one’s internal clock ‘ticks’ at a faster rate (as in instances of low inspection in sub-event monitoring), objective time appears to slow down. Conversely, if one’s internal clock ‘ticks’ at a slower rate (as in instances of close inspection in sub-event monitoring), objective time appears to speed up.

Now, arguments can be made that the above does not strictly constitute a simulation effect. If such effects can be chalked up to overlapping neurological regions and/or processing load considerations, then one needn’t postulate that simulation effects are also a driving factor. However, this argument fails to account for the entirety of the data presented in this experiment. Specifically, what the above successfully explains is how frequency modulates temporal simulation effects. Beyond this frequency-based modulation, however, it still is readily apparent that there are differences in lexical class which can affect results. Specifically, Aktionsart classes differing in their temporal density will reliably produce differing temporal simulation effects. This is captured by the two-way interaction featured above in Figure 5.12.

It was posited earlier that predication types of different temporal densities would vary with regard to how frequency modulated their elicited temporal simulation effects. In this case, it may be the case that the semantic content of low-density predication types is more readily
affected by frequency because this content, being more homogeneous and less complex, is more easily processed, streamlined, and made available for future use. Conversely, high-density predications are more stubborn: because their content is more complex and less homogeneous, it is less easily processed and streamlined for future use. Such a model is consistent with models that find that phonetic streamlining is more likely to occur on phrasal ‘chunks’ that are interpreted as single units than on otherwise identical strings that are instead interpreted compositionally (see, e.g., Bybee and Scheibman, 1999; Bybee, 2003). For example, Bybee and Scheibman found that pragmatic (i.e., non-compositional) usages of *I don’t know* were significantly more likely to undergo phonetic reduction than literal (i.e., compositional) usages of *I don’t know*, suggesting that because the latter carries with it a higher density of semantic information, it is less likely to undergo frequency-based reduction in conversation.

5.5.6 *How Lexical Classes Form*

Finally, it should be mentioned that the effect highlighted in Figure 5.12 constitutes a simulation effect that is (at least partially) a product of lexical class. In Vendler’s seminal 1957 article, Aktionsart classes were extrapolated based on their syntactic and inferential hallmarks, but these distributional patterns ultimately could not speak to whether or not Aktionsart classes were, in fact, treated as unique, discrete classes in cognitive processes.

The evidence presented here supports the claim that different Aktionsart classes are indeed processed differently in cognition (i.e., in temporal simulation). Given this, a question of primacy presents itself: it is possible that various Aktionsart classes arise as a consequence of how the human brain instinctively perceives and segments time (for general overviews of temporal processing, see, e.g., Mauk & Buonomano, 2004; Buhusi & Meck, 2005), or, alternatively, cognitive Aktionsart differentiation may arise as a result of repeated exposure to these classes of
verbs in action. For example, it may be that English speakers subconsciously learn to differentiate between stative verbs and non-stative verbs because of how stative verbs pattern in the simple present tense (e.g., they overwhelmingly tend not to occur in the present progressive while other Aktionsart classes appear much more readily in said construction). As these subconscious categories become better defined through language exposure and use, they become more richly populated with features common to all of their constituent items. In short, even though a given language user may not be able to consciously communicate the difference between the type of verb that swim is (i.e., an activity) and the type of verb that own is (i.e., a state), because swim and own (and their ilk) occur in unique environments and have unique temporal semantic features, cognitive categories differentiating between these two classes of verbs will nevertheless emerge with enough exposure (see, e.g., Slobin, 1996; Bybee, 1998; MacWhinney, 1998).

Given the above, it seems possible that because languages differ in how they lexicalize temporal semantics, cross-linguistic differences in temporal simulation may exist. We know, for example, that languages differ in the extent to which they lexicalize (or don’t lexicalize) different aspects of temporal semantics. For example, Van Valin and LaPolla (1997) (citing Everett (1986)) point out that telicity in Pirahã is added via a telicity marker rather than implicitly included in verbal semantics. To provide further evidence for this claim, it should also be pointed out that in experiment two (Chapter 4), the temporal simulation effects of temporary versus permanent states (e.g., The chef was angry versus The chef was French) were elicited from English speakers. In this case, no distinction between both types of states was found. Although it has not yet been verified, given the strongly-lexically-driven nature of the results discussed herein, it is hypothesized that a language like Spanish, which lexicalizes the distinction between
temporary and permanent states (via estar vs. ser) would show differing temporal simulation effects in this domain.

5.6 Conclusions

The predictions that were initially laid out in Section 5.1 are reiterated below:

(17) **Prediction One:** In keeping with predictions made regarding temporal density and mental simulation, predications differing in terms of their lexical aspect (i.e., Aktionsart class) will be observed to significantly differ from one another regarding the temporal simulation effects they cause. Specifically, activity predications, which are the most temporally dense of the four Aktionsart classes analyzed, are predicted to evoke relative temporal compression effects, while state predications, which have the sparsest temporal representations, are expected to evoke relative temporal dilation effects. Achievements and semelfactives are both expected to fall between these two extremes.

(18) **Prediction Two:** In keeping with the predictions set forth in sub-event monitoring hypothesis, because highly expected events should be less closely inspected than highly unexpected events, it is anticipated that predications that depict highly expected events will cause relative temporal dilation effects, while predications that depict highly unexpected events will cause relative temporal compression effects. Because closely-monitored predications are associated with a higher density of information processing, and non-closely-monitored predications are associated with a lower density of information processing, it is expected that expected and unexpected predications will be
handled in much the same way as low-temporal-density and high-temporal-density predications, respectively.

(19) Prediction Three: Once again, frequency will have a significant modulating effect on simulation results: more frequent predications will generate relative temporal dilation effects, and less frequent predications will generate relative temporal compression effects. Furthermore, it is predicted that as frequency increases, the configuration of average response times for this experiment’s various conditions (in this case, Aktionsart classes) will change (i.e., diminish) and reverse due to different degrees of sub-event monitoring among classes. It should also be noted here that this prediction may potentially countervail prediction (1), above, as the discovery of a significant interaction effect between frequency and temporal simulation effects could markedly reduce the meaningfulness of any main effects in Aktionsart class that are discovered.

Looking at the above predictions, several conclusions can be made. Firstly, predications differing in their lexical aspect were indeed observed to significantly differ in the temporal simulation effects they evoked. It should be noted that this patterning of simulation effects ran counter to predictions (i.e., stative predications were, on average, observed to generate relative temporal compression effects and activity predications were, on average, observed to generate relative temporal dilation effects). However, because Aktionsart class was shown to significantly interact with usage frequency in order to predict participants’ temporal estimation times, the fact that this finding ran counter to predictions was essentially nullified (see Section 5.5.4, above).

Secondly, a significant effect for expectedness (where more unexpected events were associated with relative temporal compression effects and/or expected events were associated
with relative temporal dilation effects) was indeed observed. This finding confirmed predictions made by the theory of sub-event monitoring, which suggested that closely-inspected predications would induce relative temporal compression effects and non-closely-inspected predications would induce relative temporal dilation effects.

Finally, following the pattern of results demonstrated in experiment two (Chapter 4), low-frequency predications were observed to significantly correlate with relative temporal compression effects, while high-frequency predications were observed to significantly correlate with relative temporal dilation effects. Crucially, it was also observed that frequency effects significantly modulated simulation-based differences between Aktionsart classes, where non-temporally-dense (i.e., state) predications were observed to be most heavily affected by frequency, and temporally-dense (i.e., activity) predications were observed to be least heavily affected by frequency.

This pattern of differential frequency effects has thus far been seen, in some form, in three different experiments in this dissertation and three different types of frequency (i.e., short-term frequency of verb-object combinations, short-term frequency of grammatical pattern features, and corpus frequency of main verbs). For instance, in this experiment, stative verbs were observed to be heavily modulated by frequency, while in experiment two, high-speed verbs were more heavily modulated by frequency affected than low-speed verbs. Given these findings, it is clear that frequency plays a heretofore undocumented but nevertheless critical role in how mental simulation plays out, and how simulation results should be interpreted. Chapter 7 will provide a thorough discussion regarding these findings.

Beyond those conclusions highlighted above, several other conclusions can be drawn here as well. Most obviously, as stated earlier, this experiment constitutes excellent evidence that temporal simulation effects are (a) real, (b) reproducible, and (c) more than just a product of
processing speed co-opting those neural regions implicated in timekeeping. This constitutes a novel finding in simulation semantics, as previous findings have primarily been limited to visual, motor, and auditory modalities. If the effects found here can be replicated further, this study offers a new paradigm for empirical inquiry within cognitive linguistics.

It also appears increasingly safe to posit that temporal density, a central construct of this dissertation, can indeed be considered a predictive factor in language-induced mental simulation. Particularly (and not surprisingly), it seems to be an important factor in the temporal simulation paradigm in linguistic items that have some sort of overt temporal semantic component. However, it should be noted that in experiment two (Chapter 4), temporal density effects were also observed across two different sets of verbs which varied in terms of their speed—not their temporal semantics. This suggests that the representational content from which temporal simulation effects emerge may not always be overtly temporal in nature, and may instead arise from the information density of the perceptual characteristics that are implicitly part of certain scenes.

Finally, as discussed above, this experiment appears to be the first simulation semantics study in which independently motivated lexical classes (i.e., Aktionsart classes) was shown to significantly modulate simulation effects. This opens up a wide array of different lexical classes across multiple languages that may be observed to generate novel and unique simulation effects.

5.7 Looking Ahead to Chapter 6

While this and previous chapters have provided novel findings about frequency, temporal density, and mental simulation, a number of research questions remain unanswered. First, while experiments two and three have consistently shown that results within the domain of temporal simulation can be modulated by long-term frequency, we do not know whether or not this type of
frequency will affect other simulation modalities, as experiment one (Chapter 3), which was a motor simulation study, only looked at short-term forms of frequency. Second, we do not know if and how multiple types of frequency may interact in mental simulation. While experiment three generated significant results for a new type of frequency (i.e., subject frequency), it was difficult to draw conclusions from this result for a number of different reasons delineated in Section 5.4, above. Third, as mentioned above, we do not know what lexical class distinctions beyond Aktionsart class-related distinctions may trigger simulation effects.

Experiment four (Chapter 6) aims to answer and resolve several of the above issues. It aims to unify two different types of frequency (i.e., short-term frequency and verbal constructional ‘preference’) in the motor facilitation and interference paradigm. This should, in theory, help to answer questions about (a) how different types of frequency may interact, and (b) how long-term forms of frequency may affect motor simulation results. Additionally, experiment four investigates the two constructions observed to occur in the so-called English dative ‘alternation’—the ditransitive and oblique goal constructions—in order to see if these constructions produce differing motor simulation results. To this end, it also looks at two independently-motivated lexical classes that are both strongly associated with the English dative opposition, Rappaport-Hovav and Levin’s (2008) verbs having only a caused possession meaning and verbs having both a caused motion and caused possession meaning. Experiment four thus also aims to show that lexical class effects in mental simulation are (a) a replicable phenomenon, and (b) observable in multiple simulation modalities.
Chapter 6: Constructional Frequency, Lexical Class and Motor Simulation Effects

6. Introduction

A central message of the work encountered thus far in this dissertation is that simulation effects are strongly driven by lexical and pragmatic factors like usage frequency and lexical class. To continue this thread of research, experiment four addresses how motor simulation effects are modulated by verbal constructional ‘preference’ and lexical class distinctions. Before directly addressing this topic, however, I will briefly review the findings of previous chapters.

Thus far, the results of experiments one, two, and three have yielded several novel findings. In all three, usage frequency, in some form, was shown to predict the distribution of the data. In experiment one (Chapter 3), short-term usage frequency (i.e., number of verb-object combination mentions) was shown to significantly modulate simulation effects, where congruent and incongruent conditions were shown to be differently affected by short-term frequency. In experiment two (Chapter 4), predications depicting situations of differing temporal density were shown to generate significantly different temporal simulation effects, and corpus frequency was also shown to significantly predict simulation effects in high-speed and low-speed verbs. Finally, in experiment three (Chapter 5), predications of different canonical ‘expectedness’ levels were shown to generate significantly different temporal simulation effects, and corpus frequency was once again shown to significantly modulate temporal simulation effects, in this case differentially affecting four distinct Aktionsart classes in English.

It should be noted here that in all three experiments, a clear pattern of differential effects of frequency upon mental simulation has also emerged. Specifically, we saw that as frequency
increases (over both the short term and the long term), differences between participant responses for various experimental conditions (e.g., congruent versus incongruent in the motor facilitation and interference paradigm, high speed verbs versus low speed verbs in the temporal estimation paradigm) diminish and then reappear in an inverted configuration. This pattern can ultimately be attributed to the different tiers of sub-event monitoring interacting with one another in specific ways (for example, frequency effects are weaker for high-temporal-density predications and stronger for low-temporal-density predications). Accordingly, there appear to exist several lexico-pragmatic factors that work in concert to determine how language-induced mental simulation will play out.

Experiment four thus has several aims. Firstly, it attempts to supplement the findings from experiment one, which demonstrated that motor simulation effects can indeed be modulated by short-term frequency. Secondly, it attempts to verify the theory built around the findings from the first three experiments (i.e., the theory of sub-event monitoring) using the motor facilitation and interference paradigm. We have yet to see a measure of long-term frequency applied to the motor facilitation and interference paradigm, and experiment four aims to address this gap using a form of collocational frequency. Specifically, experiment four takes into account English dative-alternating verbs’ constructional ‘preferences,’ (i.e., constructional biases) and investigates whether or not such preferences modulate simulation effects. Additionally, because experiment three demonstrated that lexical class (i.e., Aktionsart class) distinctions significantly modulate simulation effects, experiment four aims to follow up on this finding using lexical classes posited by Rappaport Hovav and Levin (2008) as determining the patterning of the so-called English dative ‘alternation.’ Rappaport Hovav and Levin (2008) posit that the patterning of the English dative alternation is driven, in part, by two different classes of verbs—verbs having only a change of possession reading, and verbs having both caused motion
and change of possession readings. Thus, this study distinguishes transfer verbs according to the motion/possession division, and determines whether this division has a significant influence on motor simulation effects.

Like experiment one, this experiment includes a measure of short-term frequency, as well. In this case, the aim was to verify that the findings related to short-term frequency in experiment one were indeed valid. In this case, a more limited measure of short-term frequency was employed, and this will be discussed in greater detail in 6.1.1, below.

6.1 Background Information

In this section, I discuss two areas of import to this study. First, in Section 6.1.1, I introduce the reader to the lexical classes to be featured in this study, caused possession verbs and caused motion verbs, both of which are based off of classes posited by Rappaport Hovav and Levin (2008). Following that, in Section 6.1.2, I discuss the motor simulation methodology to be used in this experiment, as it differs slightly from that used in experiment one.

6.1.1 The So-Called Dative ‘Alternation’ and Verb Classes

The English dative ‘alternation’ consists of two different constructions, the ditransitive construction and the oblique goal construction. In the active voice, the ditransitive construction is identified by a postverbal noun phrase recipient and noun phrase theme argument. The following sentence exemplifies this pattern:

(1) Lisa gave Pongo the treat.

Similarly, active-voice postverbal configurations containing a noun phrase theme argument and prepositional phrase oblique goal are categorized as being instances of the oblique goal construction:
(2) Lisa gave the treat to Pongo.

The distransitive and oblique goal constructions are often described as participating in the English dative ‘alternation,’ implying that the two constructions are equivalent to some degree, and that speakers may use either construction felicitously in any situation in which the other could be used felicitously. This implication is false, as Goldberg (1995) has argued:

[A]pproaches that rely on transformations … posit an often unwarranted asymmetry between two constructions that are thought to be related. In the case of the ditransitive, *He gave the book to her* is usually supposed to be more basic than *He gave her the book* … A typical reason given is that the verbs which allow ditransitives are a proper subset of those that allow prepositional paraphrases. However, this is not actually so: *refuse* and *deny* do not have paraphrases with *to* or *for*, and neither do many metaphorical expressions. (p. 106)

In short, while there is some distributional overlap between the two alternate constructions, there are many situations in which only one of the alternates is possible. As noted above by Goldberg, we can see instances of asymmetry between the two constructions, as exemplified by the following predicalations:

(3) The manager denied the employee a raise.
(4) *The manager denied a raise to the employee.
(5) The bartender refused him another drink.
(6) *The bartender refused another drink to him.

Thus, when we consider the dative alternation, it should be borne in mind that we are discussing two distinct constructions, rather than, say, a transformational relationship or lexical rule.

There is considerable debate concerning what factors drive the selection of one construction over the other. On one side of this debate, many scholars posit that the patterning of the dative alternation is primarily a problem of discourse pragmatics and information structure (see, e.g., Erteschik-Shir, 1979; Givón, 1984; Thompson, 1995; Wasow, 2002; Ruppenhofer, 2004). On the other side of this debate are scholars who posit lexical-semantic factors, including operations upon decomposed semantic structure, as conditioning the use of one alternate versus
the other (see, e.g., Mazurkewich & White, 1980; Pinker, 1989; Jackendoff, 1990; Groefsema, 2001; Levin, 2008; and Rappaport Hovav & Levin, 2008). Because a thorough examination of the merits of both sides of this debate would take us relatively far afield, I will simply state that a hybrid model is adopted here, wherein a combination of discourse-pragmatic factors may affect syntactic distributions that are largely determined by lexical semantics (see Rappaport Hovav and Levin (2008) for further discussion), and constructional meaning serves to enrich (or alter) the meaning provided by lexical semantics, as well.

For our purposes here, the most pertinent fact about the English dative alternation that needs to be addressed is that the various verbs that participate in the English dative alternation almost always ‘prefer’ one alternant over the other. For example, in a corpus study, Mukherjee (2005) found that the ratio of ditransitive construction appearances to oblique goal construction appearances for the verb *show* was 3.58, whereas for *send*, this same ratio was 0.97 (see Ruppenhofer (2004) for similar quantitative measures). In this study, this same ratio was calculated (via corpus study) for a number of different verbs, and these ratios will be provided in Section 6.3.2, below.

As stated above, verbal constructional preference is viewed in this experiment as a form of frequency, and expressing this preference as a ratio (rather than as a combination of raw scores) taps into a different form of frequency from that which raw scores would provide. Specifically, a verb with a ratio of 2:1 may appear in the corpus in the ditransitive and oblique goal constructions a total number of 3 times or 300,000 times. What concerns us here, however, is not how often a verb is used, so much as *how* it is used—which construction speakers are more likely to pick along with a verb in order to express an idea. Thus, a verb that has a ratio of 2:1 is more frequently used in the ditransitive construction whenever it is used, while a different verb with a ratio of 1:2 is more frequently used in the oblique goal construction whenever it is used. It
is anticipated that because this type of frequency is fundamentally different from the types of frequency used previously in this dissertation (short-term mentions, long-term usage), it may pattern differently. This possibility will be addressed in Section 6.2, below.

It is hypothesized that constructional preference should significantly predict simulation effects because constructional preference serves as a measure of the environments in which a verb frequently or infrequently occurs. In other words, a verb appearing in a frequently occurring syntactic environment should, in theory, be less closely monitored than a verb appearing in a relatively novel syntactic environment. Given this, we anticipate that as verbs’ constructional environments shift from highly unmarked (e.g., a verb that strongly favors the ditransitive construction appearing in the ditransitive construction) to highly marked (e.g., a verb that strongly favors the oblique goal construction appearing in the ditransitive construction), the character of simulation results will change such that we will again see a pattern of differential effects of frequency upon mental simulation as we did with the previous experiments seen in this dissertation.

For Rappaport Hovav and Levin (2008), a verb’s constructional preference is determined by (among other things) its lexical class. Specifically, Rappaport Hovav and Levin propose that there are two major classes of verbs that participate in the English dative alternation: (1) caused possession verbs, and (2) caused motion verbs. In the case of the former category (which includes verbs like give and sell), when these verbs are used in either the ditransitive or oblique goal constructions, they express some form of caused possession (e.g., verbs expressing physical change of possession, verbs of future having, verbs of communication, etc.). The latter category,

---

6 It should be quickly noted here that Rappaport Hovav and Levin refer to the above classes of verbs as “dative verbs having only a caused possession meaning” and “dative verbs having both caused motion and possession meanings,” as the latter class is observed to have a caused motion meaning when appearing in the oblique goal construction and a caused possession reading when appearing in the ditransitive construction. These labels are abandoned in favor of ‘caused possession verbs’ and ‘caused motion verbs’ for the sake of clarity here, even though this does, admittedly, oversimplify Rappaport Hovav and Levin’s work to some extent.
on the other hand (which includes, e.g., verbs of sending, verbs of instantaneous causation of ballistic motion, verbs of instrument of communication, etc.), expresses caused possession in the ditransitive construction, but caused motion in the oblique goal construction. Thus, we see a semantic asymmetry between the two classes of verbs. This can be illustrated, for example, with the following set of sentences:

(7) #Anna gave Debbie the present, but she never received it.
(8) #Anna gave the present to Debbie, but she never received it.
(9) #Anna sent Debbie the present, but it never got there.
(10) Anna sent the present to Debbie, but it never got there.

According to Rappaport Hovav and Levin, sentence 10 above is felicitous but 7, 8, and 9 are not because if transfer of possession is entailed either by a predication’s construction or by its main verb, then it is incoherent to deny that transfer of possession occurred. Conversely, if change of possession is not entailed by a predication’s construction or main verb, then stating that a transfer of possession did not occur is perfectly felicitous.

Given all of the above, then, we expect that caused possession verbs should produce identical simulation effects in both ditransitive and oblique goal constructions, as their meaning is posited by Rappaport Hovav and Levin to stay constant across both types. However, in the case of caused motion verbs, because this class of verbs is posited to evoke caused possession in the ditransitive construction but caused motion in the oblique goal construction, we should see simulation-based differences between these two constructions when this class of verb is utilized.

The reason why we expect to observe differences between scenes of caused motion and caused possession is because the latter necessarily involves (a) an act of transfer and (b) an animate recipient. Scenes of caused motion, on the other hand, do not necessarily involve either of the above (see e.g., Goldberg, 1995; Bergen & Chang, 2005; Goldberg, 2006; Rappaport Hovav & Levin, 2008). For example, in depicting a scene of caused motion, one could speak of...
kicking a football to the fifty-yard line, but in this case the fifty-yard line is better interpreted as a goal than as a recipient. In short, although both caused possession and caused motion scenes should evoke mental simulation, the mental simulation evoked by caused possession predications should be prototypically more likely to have a manual motor component by virtue of the fact that the presence of a recipient makes physical manual transfer a much more likely outcome.

6.1.2 The Motor Facilitation and Interference Paradigm as Instantiated in this Experiment

Experiment one (Chapter 3) looked at motor simulation effects in Spanish as modulated by short-term frequency and at a number of different grammatical patterns. As expected, short-term frequency did indeed significantly modulate simulation effects (various instances of differential frequency effects upon motor simulation results were observed), but, crucially, the simulation effects observed, on average, ran contrary to previous results within the motor facilitation and interference paradigm. Specifically, within button-is-far trials, it was observed that congruent trials generate faster response times than incongruent trials. However, within button-is-close trials, it was observed that congruent trials generated slightly slower response times than incongruent trials. It was ultimately suggested that this observed asymmetry was due to differences in habituation between close and far trials, as well as congruent and incongruent trials.

The above finding serves to highlight the variability of motor simulation effects, and it was made possible by the fact that experiment one utilized both ‘away’ verbs (e.g. throw, toss, lob, etc.) and ‘toward’ verbs (e.g., drink, take, pull, etc.). However, because experiment four is concerned with differences between the two constructions (as well as the posited lexical classes) that participate in the English dative alternation, only predications with main verbs that denote acts of transfer away from the body (with verbs like, e.g., throw, send, and hand) were implemented.
Although this design does not allow us to revisit (and, ideally, address) the pattern of results seen in experiment one, it helps to simplify the experiment’s design, thus delivering cleaner, more interpretable results. More information on this altered design is provided briefly in Section 6.3.3.

6.2 Predictions

In this section, I offer the major predictions for this experiment. They follow immediately below:

(11) **Prediction One:** In the previous experiments featured in this dissertation (experiments one through three), frequency has been shown to not only significantly modulate simulation effects as a whole (as a significant main effect), but also to significantly affect other factors modulating simulation (e.g., Aktionsart class, as part of a significant two-way interaction). Here, it is anticipated that both types of frequency investigated (i.e., short-term frequency and verbal constructional ‘preference’) will have a significant modulating effect on elicited simulation effects.

(12) **Prediction Two:** It is also anticipated that these different forms of frequency will interact with other independent variables. Specifically, it is anticipated that a verb’s constructional preference will interact with the construction in which it appears and the semantic congruity condition in which it appears such that, for example, a verb that ‘prefers’ to appear in one construction (e.g., the oblique goal construction) may generate stronger facilitation or interference effects in that construction than in its dispreferred construction (e.g., the ditransitive construction).
Prediction Three: As lexical class was demonstrated to significantly predict simulation effects in experiment three, it is predicted that simulation effects here will also be significantly predicted by Rappaport Hovav and Levin’s (2008) lexical classes. Because these classes behave differently in different syntactic environments, however, we expect there to be a three-way interaction effect between sentence congruity, constructional frame, and lexical class. Specifically, I predict that there should be no difference in motor simulation effects between caused possession verbs appearing in the ditransitive construction and caused possession verbs appearing in the oblique goal construction. As the signification in both instances is one of a possession change, it is expected that predications in both environments will be monitored similarly, thus creating similar simulation effects in both environments. In the case of caused motion verbs, however, it is expected that motor simulation differences between the ditransitive and the oblique goal construction will indeed be observed. This is specifically because in the ditransitive construction, these verbs evoke a caused possession sense, while in the oblique goal construction, these verbs evoke a caused motion sense. Because caused motion is not associated with transfer of possession or an animate recipient, the motor simulation effects evoked in instances of this sense will be weaker or entirely absent.

These predictions and other findings related to the data will be revisited and discussed in depth in Section 6.5, below.
6.3 Methods

In the following sections, I outline the methods used in this study. Specifically, Section 6.3.1 highlights those who participated in the study, Section 6.3.2 highlights the materials (i.e., stimuli) used in the study, and Section 6.3.3 highlights the study’s design and procedure.

6.3.1 Participants

Forty native English speakers over the age of eighteen years participated in the experiment for either course credit or monetary compensation. All forty participants successfully completed the experiment. These participants ranged in age from 18 to 50 years of age (average age = 23.32, SD = 7.25). Thirty-six participants reported themselves as being right-handed, two participants reported themselves as being left-handed, and two participants reported themselves as being ambidextrous. All participants had normal or corrected vision.

Participants were also asked about their linguistic backgrounds. In this case, all forty participants reported themselves as being native English speakers. Thirty-seven of these participants reported themselves as being native monolinguals, while the remaining three participants were native bilingual. Of the bilingual speakers, two spoke Spanish and English, and one spoke Korean and English.

6.3.2 Materials

56 critical stimuli were grouped with 64 non-critical stimuli to yield a total of 120 stimulus sentences. These stimulus sentences were broken up into two blocks of sixty trials each. Critical stimuli were tagged for several different variables: Dative alternation variant used (ditransitive construction or oblique goal construction), verb constructional ‘preference’ (expressed as the ratio of the number of instances in the British National Corpus where the verb appears in the
ditransitive construction to number of instances in the British National Corpus where the verb appears in the oblique goal construction), verb class according to the taxonomy provided by Rappaport Hovav and Levin (2008) (caused possession or caused motion) sentence-to-button-press congruity (congruent or incongruent), and number of mentions of the stimulus’s main verb in the experiment (1-4). Overall, fourteen different verbs were used in the study, and each verb appeared four times (twice in the ditransitive construction and twice in the oblique goal construction). Seven of these fourteen verbs were members of Rappaport Hovav and Levin’s ‘caused possession’ class, while the other seven were members of their ‘caused motion’ class.

These class memberships (along with each verb’s constructional ratio) are provided in Table 6.1 below:

<table>
<thead>
<tr>
<th>Verb</th>
<th>Class</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fling</td>
<td>Caused Motion</td>
<td>0.813</td>
</tr>
<tr>
<td>Flip</td>
<td>Caused Motion</td>
<td>0.667</td>
</tr>
<tr>
<td>Lob</td>
<td>Caused Motion</td>
<td>0.4</td>
</tr>
<tr>
<td>Send</td>
<td>Caused Motion</td>
<td>0.472</td>
</tr>
<tr>
<td>Slide</td>
<td>Caused Motion</td>
<td>0.852</td>
</tr>
<tr>
<td>Throw</td>
<td>Caused Motion</td>
<td>0.981</td>
</tr>
<tr>
<td>Toss</td>
<td>Caused Motion</td>
<td>0.545</td>
</tr>
<tr>
<td>Give</td>
<td>Caused Possession</td>
<td>2.98</td>
</tr>
<tr>
<td>Hand</td>
<td>Caused Possession</td>
<td>0.616</td>
</tr>
<tr>
<td>Lend</td>
<td>Caused Possession</td>
<td>0.498</td>
</tr>
<tr>
<td>Loan</td>
<td>Caused Possession</td>
<td>1.353</td>
</tr>
<tr>
<td>Offer</td>
<td>Caused Possession</td>
<td>1.649</td>
</tr>
<tr>
<td>Pass</td>
<td>Caused Possession</td>
<td>0.494</td>
</tr>
<tr>
<td>Show</td>
<td>Caused Possession</td>
<td>4.135</td>
</tr>
</tbody>
</table>

Although it was discussed earlier in Section 6.1.2, it should be briefly noted here again that all of the verbs featured in experiment four prototypically denote some form of manual motion away
from the body, while none of the above are readily construed as denoting some form of manual motion toward the body.

Stimulus sentences were designed with certain specifications in mind so as to maximize consistency across trials and minimize the extent to which other linguistic factors could be affecting participant responses. As such, every stimulus sentence presented to participants began with the English second person singular pronoun you. Verbs were also presented every time in the present progressive, as prior research has indicated that the imperfective aspect elicits stronger simulation effects in the motor facilitation and interference paradigm than other grammatical patterns (see, e.g., Bergen & Wheeler, 2010). Each stimulus sentence’s recipient argument was also presented as a disyllabic proper name (e.g., Ernie, Sally, Lisa), and each stimulus sentence’s theme argument was presented as a definite noun phrase. Thus, all stimulus sentences occurred in one of two general forms: (1) You are <verb>ing <name> the <noun> or (2), You are <verb>ing the <noun> to <name>.

Critical trials were generated by inserting any of the verbs featured in Table 6.1, above, into one of the above two constructions. Each verb was paired with two different theme objects such that each verb-object pairing was seen by participants in both the ditransitive and oblique goal constructions. These verb-theme pairings were determined by selecting themes that would pair naturally with their partner verbs. Thus, fling was paired with frisbee but not basket, while offer was paired with basket but not frisbee, and so forth.

Finally, filler trials were created by simply taking verbs whose argument structure configurations were incompatible with both the ditransitive and oblique goal constructions, and plugging them into the two constructional templates mentioned above. This yielded sentences like You are snoring Christian the whistle and You are existing the child to Harold. As the task required
participants to make grammaticality judgments about the sentences they saw, these stimuli served as the ungrammatical counterparts to the grammatical critical trials.

6.3.3 Design and Procedure

Participants were tested in one session that lasted roughly fifteen minutes from start to finish. Participants were seated at a desk with a laptop computer and response collection apparatus (see Section 2.3.1) and asked to fill out a brief survey detailing their linguistic and cognitive background information. Following this, the main task was commenced. Instructions asked participants to decide whether or not the sentences they saw were grammatical. Following a brief block of practice trials \(n = 5\), any questions on the part of the participant about the experiment procedure were answered, and then participants were free to complete the experiment in self-paced fashion. Halfway through the experiment, an optional break was provided to participants.

Each individual trial of the experiment began with a blank screen. The act of holding down the central yellow button on the response collection apparatus would cause a stimulus sentence to appear onscreen. As soon as the participant released the yellow button, the onscreen sentence would disappear (pressing the yellow button down again would not cause it to reappear), and participants would then press either the green button or the red button to record a judgment. After one of these response buttons had been pressed, the screen would briefly flash to let the participants know that the response had been received and that the participant had progressed to the next trial in the experiment.

This study employed a mixed design which had both within-subjects and between-subjects dependent variables. Specifically, of the variables mentioned above in Section 6.3.2, all were within-subjects except for sentence-to-button press congruity. In the case of this variable,
one half of the experiment participants (n = 20) completed the task with the button in the sentence-congruent location (i.e., farthest away from the body), while the other half (n = 20) completed the task with the button in the sentence-incongruent location (i.e., closest to the body).

The main dependent variable of this study was response time—specifically, the time between participants’ releasing of the yellow button (the button used to display sentences) and the pressing of the green response button. As with other simulation-based studies in the motor facilitation and interference paradigm, variations in these response times are presumed to signal simulation-based motor interference or facilitation (depending on how effects are construed; see Section 2.1.1).

Crucially, however, as this experiment was driven by participants’ assessments of grammaticality, only critical sentences in which participants pressed down the green button (and were expected to press down the green button) were accepted as part of the data pool. Thus, of a possible 2,240 critical trial responses, 104 (or 4.6%) were not used.

### 6.4 Results

Responses less than 200 ms or greater than 1,000 ms were removed from the data. Of the 2,135 trials collected in which participants hit the correct response key (and were expected to hit the correct response key), 0 were less than 200 ms, and 97 (4.5%) were greater than 1000 ms, leaving a total of 2,038 trials used in the statistical analysis. A mixed models analysis with crossed random effects for participant and item was then performed. This analysis yielded one significant main effect, two additional main effects that approached significance, and three significant interaction terms. Each of these findings will be briefly highlighted below, and then discussed, in turn, in Section 6.5.
Firstly, as expected, there was a significant main effect for the number of main verb mentions in this experiment (Wald $\chi^2(1, n = 40) = 101.46, p < 0.0001$), where as number of mentions increases, reaction time decreases. This effect is illustrated in the following figure (Figure 6.1):

Figure 6.1: A significant effect for short-term frequency (i.e., number of main verb mentions), where participant response time decreases as number of mentions increases. This is likely attributable to simple habituation effects.

Although the above result is indeed highly significant, it is almost certainly a product of habituation effects rather than simulation effects, as participants become more acclimated to the experimental task (and thus complete individual trials more rapidly) as experimental trials progress. This finding will be briefly revisited in Section 6.6, below.

As stated above, two additional main effects were observed to approach significance in the data. These were participant age ($\chi^2(1, n = 40) = 3.13, p = 0.08$) and sentence congruity
(congruent versus incongruent) ($\chi^2(1, n = 40) = 2.88, p = 0.09$). The relationship between participant age and reaction time is summarized in the following figure (Figure 6.2):

Figure 6.2: A direct relationship between participant age and response time that approached significance.

Above, a direct (albeit statistically insignificant) correlation between age and response time is observed, where average response time increases with age. Such a result is most likely a product of an age-related mild slowing of reflexes, and is likely not related to simulation effects.

The main effect for sentence congruity (button-press-is-congruent versus button-press-is-not-congruent) which approached significance is illustrated in the following figure (Figure 6.3):
**Figure 6.3:** Average participant response times for congruent and incongruent button press conditions.

In Figure 6.3, the relationship between congruent and incongruent conditions is illustrated. In particular, trials in which the button press location is congruent with sentence meaning (i.e., button is located farther away from the body) are observed to require, on average, 506 ms, while trials in which the button press location is incongruent with sentence meaning (i.e., button is located closer to the body) are observed to require, on average, 539 ms. Given that close button press filler trials were observed to require 569 ms on average, and far button press filler trials were observed to require 530 ms on average, we can perhaps cautiously portray the above results as a combination of a mild interference effect and a strong facilitation effect. Again, while this result is not statistically significant, several higher-order interaction terms do contain congruity as one of their items, so the fact that this result was not determined to achieve significance is not particularly problematic.

Next, a significant two-way interaction between construction type and sentence congruity was observed in the data ($\chi^2(1, n = 40) = 4.29, p < 0.05$) (Figure 6.4):
Figure 6.4: A significant two-way interaction between button press congruity and constructional type. In this case, instances of the ditransitive construction are observed to generate slightly stronger simulation effects than instances of the oblique goal construction.

In the case of the above, the incongruent condition is again observed to generate higher response times, while the congruent condition is observed to generate lower response times. However, there also appears to be a notable difference between the ditransitive construction (abbreviated above as “NP-NP”) and the oblique goal construction (abbreviated above as “NP-PP”) where stronger facilitation and interference effects are generated by the ditransitive construction. Motivations for this distribution will be provided in Section 6.5, below.

The second significant interaction term observed in the data was a significant three-way interaction between construction type, verbal constructional preference, and sentence congruity ($\chi^2(1, n = 40) = 3.85, p < 0.05$). This complicated interaction is summarized in the following figure (Figure 6.5):
**Figure 6.5:** A significant three-way interaction between construction type, sentence congruity, and verbal constructional ‘preference,’ where verbs appearing in preferred constructions are generally observed to generate stronger motor simulation effects than verbs appearing in dispreferred constructions.

The above figure contains several details of theoretical import to this study and will thus be discussed in depth in the following section.

Lastly, there was also observed to be a significant three-way interaction in the data between construction, sentence congruity, and lexical class ($\chi^2(1, n = 40) = 8.61, p < 0.01$). This finding is illustrated below (Figure 6.6):
Figure 6.6: A significant three-way interaction between construction type, sentence congruity, and lexical class, where caused possession verbs are observed to generate identical motor simulation effects irrespective of the construction in which they appear, while caused motion verbs generate strong motor simulation effects in the ditransitive construction, but weak motor simulation effects in the oblique goal construction.

A Significant Three-Way Interaction for Construction, Sentence Congruity, and Lexical Class ($p < 0.01$)

Again, as the above chart illustrates a three-way interaction, it contains several complicated relationships, each of which will be discussed in depth in section 6.5, immediately below.

6.5 General Discussion

As briefly detailed in Section 6.4, this study yielded several different significant results, many of which have ramifications for the theory of mental simulation espoused in this dissertation. Broadly, it is attested here that the character of language-induced mental simulation changes based upon the degree to which the conceptual structure of a predication is inspected in sub-event monitoring. This inspection is affected by numerous factors, including, principally, frequency and temporal density, but also other factors like markedness and expectedness.
Although, as mentioned above, the data yielded a significant main effect for short-term frequency (as well as two other main effects that approached significance), this effect was attributable to simple habituation effects, and did not significantly interact with other factors in the statistical model, unlike the short term frequency measure collected in experiment one. This discrepancy between experimental findings may be due to the fact that a greater range of short-term frequency observations \((n = 8)\) was collected in experiment one than was collected in this experiment \((n = 4)\), such that participants in this experiment weren’t exposed to enough trials for the modulating effects of short-term frequency to fully develop. In short, while there was a significant main effect for short-term frequency, it tells us very little. However, there were three interaction terms that were discovered to be significant in this experiment that are important to the general theory developed in this dissertation. These will be discussed in turn, below.

The first significant interaction term that will be discussed here is the two-way interaction between sentence congruity and construction type. This relationship is detailed in the following figure (Figure 6.4, reprinted below as Figure 6.7):
**Figure 6.7:** A significant two-way interaction between button press congruity and constructional type. In this case, instances of the ditransitive construction are observed to generate slightly stronger simulation effects than instances of the oblique goal construction.

In the above, the first and second columns represent congruent and incongruent conditions (respectively) for stimulus sentences appearing in the ditransitive construction. Correspondingly, the third and fourth columns represent congruent and incongruent conditions (respectively) for stimulus sentences appearing in the oblique goal construction. In this case, a two-way interaction is observed because the character of the relationship between congruent and incongruent stimulus sentences changes based upon the construction seen by participants. In particular, it appears that, in general, simulation effects elicited by the ditransitive construction are slightly stronger than those elicited by the oblique goal construction. This is readily observable in both congruent and incongruent conditions, as average participant response times in the ditransitive construction are faster in the congruent condition and slower in the incongruent condition.

This difference in the character and magnitude of the simulation effects generated here may, in theory, be a product of the semantics of the constructions used. Specifically, the ditransitive construction entails transfer of possession, but the oblique goal construction does not
(oblique-goal sentences may express either transfer of possession or caused motion, depending on the entailments of the verb). When language users see an instance of the ditransitive construction, they may unambiguously simulate a scene in which transfer of possession occurs, whereas language users encountering the oblique goal construction may not do so. This could hypothetically lead to more strengthened and routinized simulation effects, not unlike the long-term frequency effects seen in temporal simulation in experiments two and three (Chapters 4 and 5).

The second significant interaction term was a three-way interaction between construction type, sentence congruity, and main verb constructional ‘preference.’ This interaction is detailed in the following figure (Figure 6.5, reprinted below as Figure 6.8):

**Figure 6.8:** A significant three-way interaction between construction type, sentence congruity, and verbal constructional ‘preference,’ where verbs appearing in preferred constructions are generally observed to generate stronger motor simulation effects than verbs appearing in dispreferred constructions.

The above figure breaks down participants’ average response times in several different ways. First, the x-axis constitutes a given verb’s ratio of ditransitive construction appearances to oblique
goal construction appearances in the British National Corpus. Thus, line segments left of the “1” axis mark above should be viewed as indicating instances of verbs which ‘prefer’ to appear in the oblique goal construction, while line segments right of the “1” axis mark should be viewed as indicating verbs which ‘prefer’ to appear in the ditransitive construction.

When looking at the lines themselves inside Figure 6.8, we can see that the topmost and bottommost lines correspond to the interference and facilitation conditions (respectively) for stimulus sentences appearing in the ditransitive construction. Correspondingly, the middle two lines correspond to the interference and facilitation conditions for the oblique goal construction. Thus, when looking at the lines that correspond to the ditransitive construction, we can interpret the right side of these lines as representing verbs which prefer to appear in the ditransitive construction and are, in fact, appearing in the ditransitive construction. On the left side of these lines, then, are the verbs which prefer to appear in the oblique goal construction but are instead being presented in the ditransitive construction. The opposite is true for the two middle lines (i.e., the lines corresponding to sentences presented in the oblique goal construction). In this case, the left side of these lines can be interpreted as representing verbs which prefer to appear in the oblique goal construction and, in fact, are appearing in that construction. On the right side of these lines, on the other hand, are verbs that prefer to appear in the ditransitive construction, but are being shown in the oblique goal construction.

With the above in mind, several clear patterns can now be readily observed in the Figure 6.8. First, it is immediately clear once again that the ditransitive construction generates stronger motor simulation effects than does the oblique goal construction, but at lower ratios, it does appear that the oblique goal construction may generate slightly stronger interference effects, on average, although this is only by a relatively small margin. Furthermore, in the case of sentences appearing in the oblique goal construction, we see that congruent sentences are ostensibly not
affected by verbal constructional ‘preference’ (i.e., participant response times remain constant across all ratios), while incongruent sentences are slightly affected. Most importantly, however, clearly visible is a pattern where when a verb favors a particular construction in the English dative alternation and appears in that particular construction, it generates stronger simulation effects, on average, than when it appears in its dispreferred pattern. This is perhaps best illustrated by the configuration of lines at the far right of Figure 6.8, above. In this section of the chart, verbs appearing in their preferred construction (i.e., the ditransitive construction) generate larger differences between congruent and incongruent sentences than when they appear in their dispreferred construction (i.e., the oblique goal construction), generating smaller differences between congruent and incongruent sentences.

Given the above, it can therefore be concluded that verbs appearing in their preferred alternant in the English dative alternation generate stronger simulation effects. However, in keeping with the other findings of this dissertation, it is hypothesized here that with a greater range of dative alternation ratios, a more exaggerated pattern of differential frequency effects similar to those observed in experiments one, two, and three would emerge. As observed throughout the experimental work featured in this dissertation, the character of language-induced mental simulation appears to change significantly depending on a variety of lexico-pragmatic factors surrounding how linguistic content is monitored in simulation. In this case, constructional ‘preference’ significantly determines the extent to which motor simulation effects are observed, where smaller effects are seen when verbs appear in dispreferred constructions.

Again, it is suggested here that future studies using a greater range of constructions would likely produce a pattern more similar to those seen repeatedly throughout this dissertation.

The final significant effect observed here was a three-way interaction effect between construction type, sentence directional congruity, and lexical class. In this instance, recall
Rappaport Hovav and Levin’s lexical-class-driven theory of the dative alternation discussed above in Section 6.1.1: the authors propose that there are two major classes of verbs which participate in the English dative alternation: caused possession verbs (abbreviated as “CP,” below) and caused motion verbs (abbreviated as “CM” below). In Section 6.2, several predictions regarding the patterning of the data (given the above) were made. Firstly, it was predicted that there would be no construction-driven differences in simulation effects observed for caused possession verbs, because caused possession verbs should have the same meaning regardless of what construction they appear in. Secondly, it was predicted that there would indeed be construction-driven differences in simulation effects observed for caused motion verbs, because these verbs, according to Rappaport Hovav and Levin (2008) should change in their conveyed meaning depending on the construction in which they appear. As stated above, in the ditransitive construction, these verbs should evoke a caused possession sense, while in the oblique goal construction, a caused motion sense is evoked. Figure 6.6 (reprinted below as Figure 6.9) shows the distribution of data:
**Figure 6.9:** A significant three-way interaction between construction type, sentence congruity, and lexical class, where caused possession verbs are observed to generate identical motor simulation effects irrespective of the construction in which they appear, while caused motion verbs generate strong motor simulation effects in the ditransitive construction, but weak motor simulation effects in the oblique goal construction.

Looking at Figure 6.9, we see eight columns. Columns 1-4 represent all instances of the ditransitive construction (abbreviated as “NN”), while columns 5-8 represent all instances of the oblique goal construction (abbreviated as “NP”). Caused motion verbs are shown in columns 1, 2, 5, and 6, while caused possession verbs are shown in columns 3, 4, 7, and 8. Finally, columns 1, 3, 5, and 7 represent button-press-congruent trials, while columns 2, 4, 6, and 8 represent button-press-incongruent trials.

It was originally predicted that verbs having an intrinsic caused-possession meaning would not differ in the patterning of their motor simulation effects between the ditransitive construction and the oblique goal construction. In fact, looking at columns 3, 4, 7, and 8, we can see that they pattern virtually identically. In the case of caused possession verbs appearing in the ditransitive construction, we see average response times of 505 ms in the congruent condition,
and 537 ms in the incongruent condition. Similarly, these same verbs appearing in the oblique goal construction generate average response times of 504 ms in the congruent condition and 536 ms in the incongruent condition. In short, there is virtually no difference between these conditions, as predicted.

Turning to caused motion verbs, let us recall that there should be a significant difference between how these verbs pattern in the ditransitive construction and how they pattern in the oblique goal construction. This is because caused motion verbs are posited to evoke a caused possession sense in the former construction but a caused motion sense in the latter construction. As caused possession is more likely to involve motor simulation effects, we anticipate that motor simulation effects should therefore be the strongest in the ditransitive construction. Looking at columns 1, 2, 5, and 6, we see that this is indeed the case. Specifically, very strong simulation effects are seen in this class of verbs in the ditransitive construction, where average response times are 494 ms in the congruent condition and 545 ms in the incongruent condition. Conversely, in the oblique goal construction, the differences between congruent and incongruent conditions are the smallest of any congruent/incongruent pair in the experiment (520 ms and 540 ms, respectively).

Given the above, we can conclude that the distribution of data featured in Figure 6.9 appears to validate the intricate relationship between lexical class and construction posited by Rappaport Hovav and Levin (2008). Specifically, in the case of the ditransitive construction, caused possession is mandatorily evoked, and, as such, lexical class distinctions are overridden. However, in the case of the oblique goal construction, because the construction is unmarked for change of possession, lexical class information ultimately specifies whether or not a given predication will be simulated as involving caused possession or caused motion. Looking at the data, we see that its pattern does, in fact, conform to the above. Specifically, verbs having only a
caused possession sense are indeed observed to pattern virtually identically in both the
ditransitive and the oblique goal construction, while verbs that have both a caused possession and
a caused motion sense pattern differently depending upon the construction in which they appear.
As predicted, the strongest motor simulation effects are observed in those verb/construction
combinations where caused possession is evoked. If a caused motion sense is evoked, on the
other hand, motor simulation effects are markedly weaker (albeit still present, to some extent).

Perhaps most surprisingly, even though caused possession verbs evoke reliable motor
simulation effects in both the ditransitive and oblique goal constructions, caused motion verbs
evoke even stronger simulation effects in the ditransitive construction than caused possession
verbs do. Such a finding is, admittedly, somewhat counterintuitive. Given the rationale
presented above, a verb which evokes caused possession should be more likely to evoke motor
simulation effects, while a verb that evokes caused motion should be less likely to evoke motor
simulation effects, as caused motion doesn’t entail successful transfer or the presence of a
recipient.

A few explanations present themselves as possible solutions to the above problem. Firstly,
it could simply be the case that, regardless of lexical class or construction, all instances of caused
possession evoke relatively similar simulation patterns, and the differences seen here between
caused possession and caused motion verbs in the ditransitive construction and caused-possession
verbs in the oblique goal construction are simply a product of chance. Alternatively, it may be
that the ditransitive construction, in general, evokes stronger motor simulation effects than the
oblique goal construction does, but if a caused-possession verb is in play, these unique
constructional effects are negated, trumped, or overridden by lexical class semantics.

Experiments three and four have demonstrated that different independently motivated
lexical classes which vary in terms of their semantics and syntactic behaviors will reliably produce
different simulation effects. In experiment three, different Aktionsart classes produced significantly different temporal simulation effects, and in Experiment four, different lexical classes participating in the English dative alternation produced different motor simulation effects. What is perhaps most striking about these findings is how much these lexical classes differ in terms of their character. To wit, Aktionsart classes are a cornerstone of temporal semantic description, have clearly distinct morphosyntactic and inferential behaviors, and are found across languages. Rappaport Hovav and Levin’s verb classes, on the other hand, may be characterized as being quite the opposite. These classes are not universal, for example, and are much less well established and documented than, say, Aktionsart classes (although, to be fair, Rappaport Hovav and Levin (2008) do address their findings to a limited extent through the lens of other languages (like, e.g., Russian)).

Irrespective of how well established or paradigmatic the aforementioned lexical classes are, both Vendler (1957) and Rappaport Hovav and Levin (2008) posited these lexical classes by virtue of their distinct ways of patterning in natural language. Vendler, for example, famously suggested using verbs in certain heuristic questions to see if they could be interpreted felicitously:

\begin{align*}
(14) \text{At what time did the train arrive?} & \quad \text{(Achievement)} \\
(15) \# \text{At what time did you know Spanish?} & \quad \text{(State)}
\end{align*}

Similarly, Rappaport Hovav and Levin have suggested that combinatoric differences may reveal lexical additional class distinctions:

\begin{align*}
(16) \text{I sent the package to Sally, but she never received it.} & \quad \text{(Caused Possession or Caused Motion)} \\
(17) \# \text{I handed the package to Sally, but she never received it.} & \quad \text{(Caused Possession Only)}
\end{align*}

In short, in both of the above cases, lexical classes were posited based on certain groups of verbs’ characteristic combinatoric behaviors. This suggests that simulation-based differences between different lexical classes may, in theory, arise not just because of, say, the intrinsic semantics of a
given verb, but also because of the syntactic environments in which those lexical classes regularly appear.

In this experiment, we have seen preliminary evidence that certain syntactic environments (the ditransitive and oblique goal constructions) are associated with differing simulation effects, and we have seen evidence that lexical class semantics may negate these effects, as well (as in the case of caused possession verbs occurring in the oblique goal construction). Given this, future simulation semantics studies would do well to not only pay attention to the lexical class semantics of the different linguistic items that are tested, but also to the different types of environments in which these items appear. In short, the interplay between lexeme and construction is a complicated one, and may affect simulation in numerous ways. What is encouraging, however, is that this complex interplay provides ample opportunity for future study.

6.6 Conclusions

Returning to our predictions from Section 6.2, above, it was anticipated that (a) both short-term frequency (i.e., number of verb mentions) and long-term frequency (i.e., verb constructional ‘preference’) would significantly predict simulation effects, (b) these forms of frequency would interact with other independent variables such that we would observe that other factors predicting simulation effects (e.g., sentence-to-button-press congruity) would be modulated by frequency effects. Lastly, it was also predicted that (c) the lexical classes posited by Rappaport Hovav and Levin (2008) would also significantly predict the distribution of participants’ responses.

Given the findings discussed above, it is safe to say that all three of the above predictions were confirmed. Not only were frequency-based factors (specifically, verbal constructional
preference) significantly shown to modulate simulation effects (as well as interact with other
determining factors), but lexical classes were also shown to significantly predict simulation effects.

Nevertheless, there were some findings that did not pattern as one would expect, given the results
of previous experiments (in particular, experiment one). For instance, in this experiment,
although number of short-term mentions was indeed found to significantly predict participants’
response times, it was not, in this case, found to significantly modulate simulation effects in any
two or three-way interaction terms. This could potentially be due to the fact that in this
experiment, each verb was only seen four times, while in experiment one, each verb was seen
eight times. It stands to reason that with an insufficient quantity of exposures, short-term
frequency may not be able to come into play and thoroughly modulate simulation effects. Of
course, it should be mentioned that, given the significant (and significantly varying) habituation
effects observed in experiment one, a limited number of exposures does make for more easily
interpreted main effects in a set of motor simulation data.

Turning to the more general themes of this dissertation, we find that the predictive power
of the theory of sub-event monitoring is once again supported. Specifically, it was again shown
that frequency significantly influences simulation effects, this time in the domain of motor
simulation. While experiment one showed that number of mentions over the short-term
significantly and differentially modulated sentence congruity effects and grammatical pattern
processing time, experiment four showed that constructional preference strongly dictated the
strength of simulation effects, where when a verb appeared in its preferred construction, stronger
simulation effects were generated. This constructional preference was expressed as a ratio, a
form of type frequency that is (a) by definition, inherently relativistic, and (b) by its virtue of being
relativistic, quite different from the other forms of frequency employed in this dissertation. That
a relativistic form of frequency was also shown to significantly predict simulation effects further
serves to bolster one of the central claims of this dissertation: frequency is a critical determinant of how sub-events are monitored and inspected, and is therefore a crucial factor in how simulation plays out.

Similar to findings on frequency were findings concerning lexical class, where, in a fashion akin to the results of experiment three, stimulus sentences’ main verbs’ lexical classes were demonstrated to significantly predict simulation effects. In this case, two verbal lexical classes posited by Rappaport Hovav and Levin (2008) were tested, where the two classes were alleged to differ insofar as what semantic schemas they evoked in the two different syntactic environments of the English dative alternation. It was found that, in general, whenever a verb had a caused possession reading, motor simulation effects were quite strong, but whenever a verb had a caused motion reading, motor simulation effects were relatively weak. The strongest range of motor simulation effects was seen, surprisingly, in the oblique goal construction, which, unlike the ditransitive construction, is unmarked for a reading of caused possession or caused motion. As discussed above, this surprising finding may have simply been due to chance, or it may have been due to lexical class semantics overriding constructional semantics in instances where caused possession verbs appeared.

Finally, as briefly mentioned above, the findings of this experiment once again have fairly important ramifications for simulation semantics methodological and experimental design. Again, a measure of frequency (in this case, verbal constructional preference) was shown to have the potential to severely skew results. For instance, if one were to design a study similar to the one featured in this chapter, but use only verbs that heavily favored the oblique goal construction, then it is likely that very little difference (if any) between the oblique goal construction and the ditransitive construction would be seen insofar as motor simulation effects. Secondly, the interaction between lexical class and construction was shown to be a potential
confounding factor in simulation semantics results. Here, it was shown that the meaning of certain lexical classes appears to ‘trump’ constructional meaning, while other lexical classes cannot do the same thing. Thus, if one were to design a study similar to this one wherein the motor simulation effects of English dative-alternating verbs were tested in both forms of the English dative alternation, differentiating effects could be nullified or diminished if one failed to pay sufficient attention to the specific combination of verb class and constructional frame in each individual stimulus sentence.

### 6.7 Looking Ahead to Chapter 7

Although experiment four is the final experiment featured in this dissertation, once again, its findings raise several possibilities for future research. These will be briefly discussed here. Based on the discussion provided in Section 6.6, the most obvious path for future research would be to continue to investigate the relationship between syntactic environment and lexical class in determining simulation effects. This could be done by looking at other English syntactic alternations (e.g., the ‘spray/load’ alternation, the verb-particle construction alternation) in order to see if significantly differing simulation effects could be observed in these domains, as well. Furthermore, future studies could be performed that (a) looked at pure constructional meaning (via the use of nonce words, as in, for example) Sue vorped Bill the dax, or (b) coerced meaning (as in, for example, Ron sneezed Debbie the napkin). Finally, and perhaps most obviously, a future dative alternation study with a wider range of constructional ‘preference’ ratios would be ideal in order to see if the more extreme differential frequency effects observed in other experiments would indeed be observed in this instance, as well.

The following chapter (Chapter 7) is the final chapter of this dissertation. In it, a review of experimental findings will be provided, along with an extended discussion of how these
findings may be integrated into the theory of sub-event monitoring, and, of course, into a larger theory of simulation-based representation.
Chapter 7:
General Discussion and Conclusion

7. Simulation Semantics through the Lens of Cognitive-Functionalist Linguistics (and Vice-Versa)

Simulation semantics has helped to challenge predominant views in cognitive science, psychology, and linguistics, and has ultimately changed how we conceptualize linguistic processing and representation in the brain. However, the approach remains problematic for a number of reasons. For one, as discussed in Chapter 2, simulation semantics studies often have inconsistent and highly variable results—particularly within the motor paradigm. Moreover, the range of response times reported may change considerably depending on a number of different factors, including methodological variations on the experimental paradigm itself (see Section 2.3.1 for more detail).

Another major issue with simulation semantics is that simulation, regarding its cognitive instantiation, has been viewed as a relatively homogenous phenomenon. Although Barsalou (1999) points out that (a) simulations can be used for a wide variety of cognitive functions (including representing concrete scenes or abstract concepts), and (b) simulations affect domains including language, working memory, long-term memory, and problem solving, experimental research within simulation semantics tends to treat simulations as simple facilitation and/or interference effects. In reality, however, the fundamental character of simulations may change from situation to situation and predication to predication.

The work done within simulation semantics can likely be characterized as problematically homogenous in other ways, as well. Specifically, the majority of simulation semantics studies are (a) performed only in English, and (b) performed within either the visual or motor modalities. In
addition, there has been very little evidence presented which supports Barsalou’s (1999) contention that simulations can be used to represent abstract meaning. It is relatively uncontroversial that simulation effects are generated when participants read or hear sentences dealing with embodied, physical action, but major discrepancies begin to appear in the literature when abstract meaning enters the picture (cf. Glenberg & Kaschak, 2002; Bergen et al, 2007). Many therefore view simulation-based models of cognition to be incomplete, as linguistic competence relies heavily upon speakers both wielding and comprehending abstraction in everyday language use (see, e.g., Zwaan, 2014).

It was hypothesized in Chapter 1 that accounting for lexico-pragmatic factors in simulation semantics experiments could help to resolve some of the issues mentioned above. In particular, it was posited that lexico-pragmatic factors (like, e.g., usage frequency) would not only be shown to significantly modulate simulation effects, but, in doing so, could also potentially help to explain the variability often seen in within and between different simulation semantics studies. Furthermore, such findings would help to elucidate the true character of mental simulations: they are not immutable, recurring patterns of neurological activation, but are instead highly variable experiential gestalts that are affected by many different factors. In the following section, specific goals surrounding the above hypothesis are revisited and reviewed in light of the findings of the four experiments presented in this dissertation.

7.1 Revisiting the Goals of the Dissertation

In Chapter 1, the goals of the thesis were provided to the reader. I repeat these goals immediately below and then discuss the extent to which each was met.
(1) Establish a link between certain lexico-pragmatic factors (in particular, usage frequency and lexical class) and simulation-based representation.

(2) Show that this link exists in multiple simulation paradigms (motor simulation and temporal simulation) and in multiple languages (English and Spanish).

(3) Conduct follow-up experiments in both paradigms in order to verify the findings in 1 and 2, above.

(4) Finally, develop a theory (sub-event monitoring) that successfully captures the patterning of all of the data gathered and that, ideally, will be able to predict the distribution of data in future experiments.

Goal 1, above, constituted the broadest of the four major goals featured in this dissertation. Here, the aim was simply to prove that lexico-pragmatic considerations could affect simulation results. In light of the results of experiments one through four, which firmly showed that such factors could enhance, nullify, or even potentially reverse the configuration of average participant response times in simulation studies, it appears safe to conclude that lexico-pragmatic factors do indeed affect simulation-based representation. Specifically, both usage frequency and lexical class were shown to significantly predict simulation results. In particular, several different types of usage frequency achieved statistical significance either as main effects or within interaction terms. These included number of mentions of verb-object combinations (experiments one and four), number of mentions of the grammatical patterns in which they appear (experiment one), main verb corpus frequency (experiments two and three), and constructional ‘preference’
Lexical classes shown to significantly predict simulation effects were Aktionsart class and Rappaport Hovav and Levin’s (2008) verbs having only a caused-possession meaning (i.e., “caused possession verbs”) and verbs having both a caused-possession and caused-motion meaning (i.e., “caused motion verbs”). It bears repeating, too, that the above lexical classes were independently motivated. In other words, unlike previous simulation semantics studies which have generated critical items via an ad hoc selection process and subsequent norming study, the lexical classes mentioned directly above were gleaned through careful observation of distributional patterns surrounding certain lexemes in spoken and written language.

In short, the first goal was met multiple times over this dissertation. Furthermore, the extent of the effects that lexico-pragmatic factors were observed to have on mental simulation effects was considerable. For instance, at least one instance of frequency differentially modulating simulation effects generated by different linguistic categories was observed in all four experiments. This important recurring pattern will be discussed in Section 7.3, below.

The second and third goals, above, represent two different means of establishing the validity of the finding that lexico-pragmatic factors do indeed affect mental simulation. Specifically, both goals are concerned with replication of findings, where Goal 2 is concerned with across-paradigm replication, and Goal 3 is concerned with within-paradigm replication. Although in this dissertation no single experiment was replicated in exact one-to-one fashion, each of the two major methodologies (the motor facilitation and interference paradigm and the temporal estimation paradigm) was repeated once, and in both cases, significant effects were observed. However, it should again be noted here that in both of these cases the material tested differed and, in the case of experiments one and four, different languages were used (Spanish and English, respectively). Because the predications used in simulations semantics studies are generally modality-specific (i.e., a critical trial in a visual simulation study would convey an
unambiguously visual scene, while a critical trial in a motor simulation study would convey a scene that unambiguously involves some sort of motor action), it becomes nearly impossible to do a perfect cross-modal replication of simulation content. Furthermore, although lexico-pragmatic factors were shown to significantly predict simulation results in all four experiments, these factors were often not the same. There was not, for example, a temporal simulation study that looked at short-term frequency effects, even though both motor simulation studies (experiments one and four) did so. Similarly, experiment four (the second motor simulation study) was the only study to look at effects of constructional ‘preference’ on simulation. In short, although this dissertation generated a rich variety of results demonstrating that lexico-pragmatic factors do indeed affect mental simulation, it nevertheless remains a distinct possibility that certain factors may modulate one type of simulation more or less strongly than they would another type of simulation.

However, given that these factors manifest themselves in language in myriad ways and forms, it is perhaps safest to assume that their effects on mental simulation are not uniform, and that future studies will be needed in order to better understand their behavior.

In light of the above, one can say that goals two and three, above, have been met, but only in the broadest possible sense. While similar lexical class effects and similar frequency effects were recorded, direct replications of these effects were not possible. Nevertheless, the relatively high number of significant lexico-pragmatic effects recorded in this dissertation does strongly bolster the claim that these factors play a large part in determining how mental simulation manifests itself.

The final major goal of this dissertation was to develop a functional theory that would not only capture the behavior of the data of the four experiments featured herein, but also serve to predict future studies’ results. This theory, called sub-event monitoring, suggested that speakers will ‘inspect’ simulation-based content more closely or less closely depending on a variety of lexical
and pragmatic factors. The following section revisits the major findings of this dissertation and attempts to integrate them with this theory.

7.2 Mental Simulation and Sub-Event Monitoring

Introduced in Chapter 2, the theory of sub-event monitoring provides a theoretical framework that links factors central to cognitive-functional linguistics (i.e., lexico-pragmatic considerations) with observed simulation effects. In general, high-monitoring predication types and low-monitoring predication types are handled quite differently in simulation due to a complex interplay between three different ‘tiers’ of content, which are summarized immediately below:

(1) Tier one consists of the basic conceptual content to be simulated found in a class of words or phrases. A commonality across words that involve manual causation of ballistic motion, for example, is that regardless of how often each one is used in everyday language, all of them contain, at their core, an experiential amalgam that activates a distributed neural representation involving manual motor action moving away from the body. The character of this basic pattern of activation is, in turn, governed by tiers two and three, which are described below:

(2) Tier two consists of usage frequency. Low-frequency (i.e., high-monitoring) words and phrases are associated with a higher processing load (yielding, e.g., temporal compression effects), while high-frequency (i.e., low-monitoring) words and phrases are associated with a lower processing load (yielding, e.g., temporal dilation effects). For example, two words or phrases with similar tier one content but of differing frequency (e.g., throw and fling)
would be inspected to different degrees, processed at different rates, and therefore yield differing simulation effects.

(3) Tier three consists of additional lexico-pragmatic factors like temporal density, markedness, lexical class, etc., which modulate the strength of frequency effects in tier two. In this case, the observed simulation effects for predication types with lexico-pragmatic factors that are associated with a low degree of inspection in sub-event monitoring will be heavily modulated by frequency, while the observed simulation effects for predication types with lexico-pragmatic factors that are associated with a high degree of inspection in sub-event monitoring will be weakly modulated by frequency. As discussed in Section 5.5.5, this may be because the semantic content of low-inspection predication types is more readily affected by frequency because this content, being more routine and less complex, is more easily processed, streamlined, and made available for future use. Because the content of high-inspection predications is more complex and less homogeneous, however, it is less easily processed and streamlined for future use.

Below, I assess whether or not the above model is able to account for the data gathered in the four experiments featured in this dissertation. It will ultimately be demonstrated that the majority of the findings encountered here can be felicitously subsumed under the theory of sub-event monitoring, although there is one specific finding from this dissertation that proves problematic for it.

Beginning with motor simulation results, the following figure represents the pertinent motor simulation effects from experiment four. In it, a pattern of differential frequency effects is readily observable (Figure 7.1):
Figure 7.1: A three-way interaction between construction type, sentence congruity, and verbal constructional ‘preference’ (as instantiated by BNC corpus counts). This chart was featured in experiment four (Chapter 6).

In the case of the above figure, a clear interplay between constructional markedness and frequency is observed. Specifically, between the ditransitive and oblique goal constructions, the oblique goal construction is unmarked (in that it may evoke either caused motion or change of possession) and is therefore monitored more closely due to its underspecification, while the ditransitive construction is marked and is therefore monitored less closely (as it only entails change of possession). Close monitoring is associated with diminished frequency effects, meaning that, in this case, as we transition in the oblique goal construction from verbs which infrequently appear in the construction to verbs which frequently appear in the construction (moving from right to left in Figure 7.1, above), the differences between these frequent and infrequent verbs are much less apparent than when we make this same transition in the ditransitive construction (moving from left to right in Figure 7.1, above). In other words, frequency effects are greater in the ditransitive construction and lesser in the oblique goal construction, and, indeed, the former is monitored less closely and the latter more closely.
As another example illustrating the interplay between sub-event monitoring and motor simulation, consider the following significant three-way interaction from experiment four (Figure 7.2):

**Figure 7.2:** Another significant three-way interaction from experiment four (Chapter 6). In this instance, construction type, sentence congruity, and lexical class are shown to interact such that caused motion verbs appearing in the ditransitive construction generate relatively strong motor simulation effects, while caused motion verbs appearing in the oblique goal construction generate relatively weak motor simulation effects. Caused possession verbs generate nearly identical motor simulation effects in both constructions, as their meaning is invariant across both types.

Above, we again see several different features associated with different tiers of sub-event monitoring, although it should be noted that this presents a unique situation in that frequency is not a factor in this chart. Nevertheless, these findings can readily be subsumed within the theory of sub-event monitoring. Comparing columns one and two (caused motion verbs appearing in the ditransitive construction (which elicits only the caused-possession sense)) with columns five and six (caused motion verbs appearing in the oblique goal construction (which in this case elicits a caused possession reading)), two different features associated with different degrees of inspection
in sub-event monitoring are in play. These are, (1) constructional markedness, and (2) lexical class markedness. The interplay between these two types of markedness is described in Table 7.1, below:

**Table 7.1:** A summary of the interplay between markedness effects in sub-event monitoring observed in Figure 7.2, above.

<table>
<thead>
<tr>
<th></th>
<th>Unmarked Class (Caused Motion)</th>
<th>Marked Class (Caused Possession)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Unmarked</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Construction</td>
<td>Smallest Effects (Columns 5 and 6)</td>
<td>Medium Effects (Columns 7 and 8)</td>
</tr>
<tr>
<td>(Oblique Goal)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Marked</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Construction</td>
<td>Largest Effects (Columns 1 and 2)</td>
<td>Medium Effects (Columns 3 and 4)</td>
</tr>
<tr>
<td>(Ditransitive)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In the case of the interplay between types of markedness presented in the table above, it appears that, insofar as simulation effects, lexical class markedness effects may supersede constructional markedness effects. Essentially, when a marked (i.e., semantically unambiguous) lexical class is involved in a simulation, the construction in which it appears has no measurable effect on simulation. However, if an unmarked (i.e., semantically ambiguous) lexical class is involved in simulation, then constructional meaning becomes a very important factor. In this case, constructions marked for transfer of possession are associated with strong motor simulation effects, while constructions unmarked for transfer of possession (i.e., those which Rappaport Hovav and Levin (2008) suggest are associated with either caused motion or caused of possession) generate relatively weak motor simulation effects.

Within the temporal estimation paradigm, complex interactions between different tiers of sub-event monitoring can be seen, as well. Firstly, looking at simple results from experiments two and three, one can see that high-monitoring predications are associated with relative temporal compression effects and low-monitoring predications are associated with relative temporal dilation effects. The following two figures illustrate this (Figure 7.3 and Figure 7.4):
**Figure 7.3:** A significant finding from experiment two (Chapter 4) where more closely inspected predications (i.e., predications of higher temporal density) are observed to generate relative temporal compression effects and less closely inspected predications (i.e., predications of lower temporal density) are observed to generate relative temporal dilation effects.

**Figure 7.4:** A significant finding from experiment three (Chapter 5) where more closely inspected predications (i.e., ‘unexpected’ predications) are observed to generate relative temporal compression effects and less closely inspected predications (i.e., ‘expected’ predications) are observed to generate relative temporal dilation effects.
In Figures 7.3 and 7.4, above, predications of differing degrees of inspection in sub-event monitoring are shown to produce significantly different temporal simulation effects. In particular, low-density or expected (i.e., low-inspection) predications are associated with temporal dilation effects, while high-density or unexpected (i.e., high-inspection) predications are associated with temporal compression effects. As discussed in Chapter 5, in particular, it is hypothesized that speed of linguistic processing affects the speed of temporal processing, as (a) temporal processing occurs in diverse and distributed neural regions (Lewis & Miall, 2003; Ivry & Spencer, 2004; Mauk & Buonomano, 2004) and (b) these regions of temporal processing overlap with regions of linguistic processing (including, perhaps most prominently, the basal ganglia) (Wallesch et al, 1983; Crosson et al, 2003, Booth, Wood, Lu, Houk, & Bitan, 2007; Crosson, Benjamin, and Levy, 2007).

As with the motor simulation results discussed above, once multiple features associated with different degrees of sub-event monitoring begin to interact, the picture becomes much more complex and interesting. The following figure (Figure 7.5) illustrates the interplay between temporal density and usage frequency discovered in experiment three:
**Figure 7.5:** A significant two-way interaction between Aktionsart class and verb frequency featured in experiment three (Chapter 5). In this case, the most temporally dense (and thus, most closely inspected) Aktionsart class (activities) is least susceptible to frequency effects, while the least temporally dense (and thus, least closely inspected) Aktionsart class (states) is most susceptible to frequency effects.

In the above figure, a clear pattern of differential effects of frequency upon temporal simulation can be seen. In particular, as discussed above, those predication types which are most closely inspected (i.e., activities) are less susceptible to frequency effects than those predication types which are least closely inspected (i.e., states).

The following table (Table 7.2) summarizes the various highlighted findings of all four experiments and shows how they relate to sub-event monitoring. Specifically, it highlights whether or not this dissertation’s major findings dovetail with the two major predictions generated by sub-event monitoring: (1) high-monitoring predications are associated with temporal compression effects while low-monitoring predications are associated with temporal dilation effects, and (2) high-monitoring predications are less susceptible to frequency effects, while low-monitoring predications are more susceptible to frequency effects. Additionally,
whether or not particular findings can be felicitously subsumed within the theory of sub-event monitoring is featured below in the rightmost column.

**Table 7.2**: A table summarizing the results of each of the experiments featured in this dissertation as they relate to the theory of sub-event monitoring.

<table>
<thead>
<tr>
<th>Experiment Number/Modality</th>
<th>Area of Focus</th>
<th>Is high monitoring associated with relative temporal compression effects?</th>
<th>Are high-monitoring predications less susceptible to frequency effects?</th>
<th>Can this finding be subsumed within the theory of sub-event monitoring?</th>
</tr>
</thead>
<tbody>
<tr>
<td>One/Motor</td>
<td>Toward/Away Verbs</td>
<td>n/a</td>
<td>Uncertain</td>
<td>Uncertain</td>
</tr>
<tr>
<td>One/Motor</td>
<td>Grammatical Pattern</td>
<td>n/a</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Two/Temporal</td>
<td>High/Low Speed Verbs</td>
<td>Uncertain</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Two/Temporal</td>
<td>Habitual/Single Events</td>
<td>Yes</td>
<td>n/a</td>
<td>Yes</td>
</tr>
<tr>
<td>Three/Temporal</td>
<td>Expected/Unexpected Events</td>
<td>Yes</td>
<td>n/a</td>
<td>Yes</td>
</tr>
<tr>
<td>Three/Temporal</td>
<td>Aktionsart Class</td>
<td>Uncertain</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Four/Motor</td>
<td>Lexical Class</td>
<td>n/a</td>
<td>n/a</td>
<td>Yes</td>
</tr>
<tr>
<td>Four/Motor</td>
<td>Constructional Markedness</td>
<td>n/a</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Above, the first aspect of sub-event monitoring (namely, that high monitoring is associated with temporal compression effects while low monitoring is associated with temporal dilation effects) appears to be moderately well supported, where of the two temporal results which are not modulated (and therefore confounded) by frequency effects, both fit our predictions. The second aspect of sub-event monitoring (namely, that high monitoring predication types are less susceptible to frequency effects while low monitoring predication types are more susceptible to frequency effects) appears to be somewhat more solidly supported. In this case, of the five different instances where frequency could modulate simulation effects, three conform to the model, while one does not, and another is inconclusive. Finally, as to whether or not the findings
highlighted above can be subsumed within the theory, six of eight are compatible, one is uncertain, and one (i.e., the modulating effects of frequency on high speed verbs and low speed verbs) is not compatible. Overall, therefore, the theory of sub-event monitoring appears to be on the right track, but, admittedly, doesn’t perfectly capture the data. In Section 7.5, below, several ideas for future studies are provided, many of which aim to directly address the various inconsistencies seen above.

7.3 On Differential Frequency Effects

As briefly noted in the section above (and in previous sections), in looking at interactions between linguistic features and usage frequency (in its various forms) in this dissertation, a common pattern appears to have emerged across all four of the experiments. Specifically, it appears that different linguistic categories may have different frequency profiles depending upon certain aspects of their linguistic content (e.g., temporal density, constructional markedness, etc.). These differential frequency effects are presented, in abstract form, in the following figure (Figure 7.6):
In the above, a number of differential frequency effects are readily apparent. In experiments one and three, the range of extremes in frequency effects was readily apparent, while in experiment two, a similar pattern was observed, although it was not viewed as being conclusive, given the range of frequencies tested. Experiment four presented a similar but more moderate instance of differential frequency effects within the motor simulation paradigm.

It was posited in the previous section that this pattern emerges due to interactions between different tiers of sub-event monitoring content, and there is indeed some evidence that appears to support this. Experiments three and four, in particular, conform to the model very well, while experiment one conforms to a more limited extent. Experiment two, on the other
hand, while displaying a pattern of differential frequency effects, actually runs counter to the predictions of sub-event monitoring in that high-monitoring predications (i.e., high-speed predications) were seen to be more susceptible to frequency effects than low-monitoring predications (i.e., low-speed predications). Now, there are many explanations that could be given as to why this result runs counter to predictions (e.g., the range of frequencies sampled was unbalanced and/or too small), but in the absence of further empirical evidence to confirm (or disprove) the validity of this result, such explanations remain little more than weak conjecture.

Returning to the general distribution of the data, irrespective of the fact that experiment two did not conform to the predictions of sub-event monitoring, the repeated patterns of differential frequency effects seen throughout the data nevertheless have important ramifications for how simulation is conceptualized. Most obviously, these results demonstrate the need for usage frequency to be taken heavily into consideration in future simulation semantics studies. In particular, they strongly suggest that the predominant method of using stimuli generated via norming study without regard for frequency may distort how data are interpreted. In this case, the configuration of participant responses could be altered or potentially even reversed if the stimuli from which they are generated aren’t controlled for frequency. Moreover, even if, for example, stimuli which all cluster around a particular frequency are used in a given experiment, there is no guarantee that a significant effect will be generated, as these stimuli could have been gathered by chance from a locus of convergence of average participant responses where little difference between conditions would be found. Furthermore, in the event that a significant result were found using this methodology, this would still not tell the entire story about the simulation effects documented, as the differential effects of frequency on simulation could cause the configuration of results to be wildly different at other frequencies. In short, simulation effects, in the future, need to be conceptualized, when possible, as interactions between first-tier simulation
content and frequency. To refer to an effect as a simple case of facilitation or interference no longer appears to be sufficient.

The differential frequency effects seen in the data also make it difficult to conceive of mental simulation as a unitary phenomenon. Rather, it appears to substantially vary in character depending on the familiarity (or lack thereof) and complexity (or lack thereof) of the linguistic content being simulated. It may be reasonable to assume that high-monitoring simulations are best utilized when one is confronted with the novel and unfamiliar. In processing this novel information, there may be little room for error and, as such, extra time needs to be spent processing representational content in order to ensure that it is used properly. Conversely, in processing familiar information, such monitoring should, in theory, not be necessary. It may therefore also be reasonable to assume that predications of identical semantic content could heavily change in their elicited simulation effects depending on the context in which they are presented. More on this notion will be presented in Section 7.4, immediately below.

7.4 General Conclusions

The research presented in this dissertation has yielded a number of findings, some of which may simply serve as interesting details to be remembered, while others may be viewed as having fairly weighty ramifications for theories of mental simulation. In the following paragraphs, these findings are presented, with less significant conclusions being followed by more significant ones.

Firstly, experiments one and four have provided evidence that grammatical patterns and constructions do indeed affect how mental simulation of sentential content emerges and plays out. Previous research (e.g., Madden & Zwaan, 2003; Madden & Therriault, 2009; Bergen & Wheeler, 2010) has shown that different aspectual constructions play an important role in how
simulation is construed, and here it was similarly shown that various tense/aspect/modality features play a role in determining the character of mental simulation. In experiment one (Chapter 3), for instance, there were significant differences between the Spanish morphological future, and its present, present progressive, and preterite. There was a significant two-way interaction between grammatical pattern and mention number, as well, where the various patterns seen in experiment one were seen to follow the trend of differential frequency effects repeatedly encountered throughout this thesis. Furthermore, in experiment four (Chapter 6), two different grammatical constructions were observed to yield significantly different motor simulation results. These were the ditransitive construction and the oblique goal construction.

Secondly, experiment two showed that in one critical contrast (the habitual event versus single event contrast), a significant difference between conditions was generated by sentences which only differed in terms of one adverb (once versus often). While previous simulation semantics studies have found that simulation effects could be caused by main verbs, subject nouns, and object nouns, this study provides the first evidence (to my knowledge) that the semantic content of a temporal adjunct could generate significantly differing simulation effects.

Thirdly, as discussed in Section 7.2 above, sub-event monitoring, the framework developed in order to bridge the gap between lexico-pragmatic features and simulation semantics, helps to explain an array of otherwise disparate patterns. While it does not perfectly account for all of the data explored here, it does, for now, serve as a simple mechanism for explaining relatively complex phenomena in simulation, like, say, interactions between frequency and temporal density, or frequency and markedness. With further refinement and additional experiments used to test the model, sub-event monitoring could develop into a powerful explanatory mechanism linking lexico-pragmatic features with simulation semantic representation.
Turning to larger conclusions, I offer four lessons from this dissertation. Each is discussed in turn below.

(5) **Major Conclusion One:** Temporal simulation appears to be a viable new simulation semantics methodology.

Two different experiments in this dissertation (experiments two and three) found significant temporal simulation effects driven by a variety of semantic and lexico-pragmatic factors. This thesis constitutes the first study, to my knowledge, to document such effects, and it therefore marks the inception of a new research paradigm wherein any number of linguistic items may be investigated, presuming they have some sort of temporal component. Some possibilities for future studies within this paradigm are listed in Section 7.5, below.

(6) **Major Conclusion Two:** Independently motivated lexical classes appear to significantly predict mental simulation.

Experiments three and four each found that different lexical classes were significantly correlated with differing simulation effects. Crucially, these lexical classes were (a) independently motivated (unlike the classes used in previous studies), and (b) observed in two different simulation modalities (i.e., temporal and motor modalities). In experiment three, four Aktionsart classes were investigated, while in experiment four, Rappaport-Hovav and Levin’s (2008) dative-alternating verb classes were investigated. In either case, the finding that class-level (rather than word-level) semantics were shown to correlate with differing simulation effects again suggests that mental simulation is the product of complex processing operations synthesizing a wide variety of linguistic and supralinguistic factors.

(7) **Major Conclusion Three:** Frequency effects and simulation-based representation are inextricable.

As stated earlier, all four experiments featured in this dissertation generated significant effects for frequency. Furthermore, the majority of these frequency effects were present in two-way or
three-way interactions with different sub-event monitoring factors (e.g., lexical class, temporal density, etc.), meaning that the effects of frequency on simulation, in this case, go beyond simple processing considerations. Rather, the different lexico-pragmatic factors implicated in sub-event monitoring appear to modulate the effects of frequency on simulation.

Many different types of frequency were featured in this dissertation, as well. This included short-term frequency (as instantiated by verb-object combination mentions in experiment one, as well as by verb mentions in experiment four), and long-term frequency (as instantiated by main verb frequency in experiments two and three, phrase-level frequency in experiment three, and verbal constructional ‘preference’ in experiment four). Other types of frequency (e.g., subject frequency) were investigated, as well, although the validity of these effects was somewhat more difficult to establish.

The significance of the role of frequency in simulation semantics cannot be overstated. It has the power to nullify, distort, or potentially even reverse the configuration of participant responses in a simulation-based study. Consequently, frequency should certainly be included as a predictor variable and an interaction term in all future simulation semantics studies.

(8) **Major Conclusion Four:** Lexico-pragmatic factors affect mental simulation, and do so in a significant way.

In order to successfully model linguistic processing of any type, we must include lexico-pragmatic considerations, or else the thing being modeled quite simply isn’t a valid reflection of natural language. Of course, lexico-pragmatic considerations are multitudinous and exceedingly complex. In fact, the work featured here has arguably not delved beyond the most superficial factors within cognitive-functionalist linguistics (i.e., frequency, lexical class, and markedness). Nevertheless, it counts for something: these factors appear to have a major effect on how mental
simulation plays out. The task of disentangling and delimiting the simulation-based effects associated with these factors is what lies ahead.

### 7.5 Possibilities for Future Study

The work featured in this dissertation has, in the end, raised more questions than it has answered, and perhaps that is a good thing, as several possibilities for future study immediately present themselves.

First, most obviously, it is clear that frequency and simulation are highly interrelated, and the research performed here has only begun to scratch the surface of this complex relationship. Here, token frequency was predominantly investigated (in the case of experiment four, it is difficult to classify a verb’s constructional ‘preference’ as either type or token frequency). Future studies may benefit from investigating the role of type frequency in mental simulation—especially those studies which look at the effects of syntax on simulation. Certain syntactic configurations may be less standard and therefore inspected more closely than other, more canonical configurations. Therefore, it may be possible to create differing simulation effects with the same semantic content, but in different syntactic environments (e.g., active voice versus passive voice).

Also with regard to frequency, although this dissertation did indeed look at the frequencies of items that were not verbs (e.g., subjects, objects, etc.), valid, significant results were not readily observed in these domains. It is difficult to take this finding as conclusive, however, because these frequency counts failed to capture the relationship between word-level frequency and phrase-level frequency (save for the ‘expectedness’ trials in experiment three). Instead of looking at individual verbs and individual objects, it may be much more telling to look at, say, the long-term frequency of certain verb-object combinations or certain subject-verb combinations. If
language is better characterized as being composed of phrasal ‘chunks’ instead of individual words, then perhaps studies addressing these forms of frequency may be more informative.

Second, the theory of sub-event monitoring presented here was able to account for a good amount of the data, but it needs to be explicitly tested and used in more experiments in order to verify its accuracy and efficacy. Doing so would entail running more studies that look for interactions between different features of sub-event monitoring and different types of frequency. For example, markedness was only briefly mentioned here in experiment four, but it is undoubtedly a central concept in modern linguistic description. If unmarked linguistic items are underspecified and must therefore be monitored more closely, then there are numerous pairs of lexemes, constructions, etc., which can feasibly be studied using simulation semantics methodology with this distinction in mind.

Thirdly, a fruitful area for future study concerns linguistic relativity effects. This is especially true for studies employing the temporal estimation methodology, as this methodology has not yet been used beyond what has been seen in this dissertation. In experiment two, for instance, English speakers were given sentences denoting both temporary and permanent states (e.g., The chef was angry versus The chef was French). In this case, however, no significant difference between these two sentence types was observed. English has no lexicalized distinction between temporary states and permanent states, while many other languages do—in particular, Spanish. It may therefore be the case that a sentence like la mujer está enferma (the woman is sick) would contrast with la mujer es enferma (the woman is sickly) in terms of temporal simulation effects in Spanish, while the English equivalents of these sentences would cause no such effects.

Another domain of relativity-based temporal study that may be fruitful in the future involves metaphor. Experiment two demonstrated that differing verb speeds were linked to differing temporal simulation effects, and many scholars have suggested that physical speed and
time are linked via metaphor (Lakoff & Johnson, 1980; Lakoff 1993; Boroditsky, 2000; Matlock, Ramscar, & Boroditsky, 2005; Casasanto & Boroditsky, 2008; Ramscar, Matlock, & Dye, 2010). Of course, it is well-known other languages have different metaphors for time and space, and it may be the case that, to take a page from some of the work of Boroditsky (Boroditsky, 2001; Boroditsky, Fuhrman, & McCormick, 2011), sentences denoting horizontal motion may induce stronger temporal simulation effects in English speakers than in Mandarin speakers (who are more likely to employ a different spatial metaphor for the passage of time), while sentences denoting vertical motion may induce stronger temporal simulation effects in Mandarin speakers than in English speakers.

Finally, given the strong effects that lexico-pragmatic factors were shown to have on mental simulation, it is clear that the role of supralinguistic factors on mental simulation needs to be investigated further. Sub-event monitoring suggests that people simulate differently depending on a number of different factors, where sometimes simulation is an intensive, detailed process, and other times it is a rapid, automatic process. It may therefore be possible to run a between-subjects study where, for example, one group of subjects is asked to read sentences in a high-pressure situation (e.g., under a time limit and knowing they’ll be quizzed on these sentences), while another group of subjects is asked to read the same sentences in a low-pressure situation (e.g., no time limits, no quiz questions). In this case, then, the type of sub-event monitoring performed by participants would not be determined by linguistic factors, but would instead be imposed by extrinsic pressures on the participant. In this case, we would ideally expect to observe that participants in the high-pressure condition would generate simulation effects patterning opposite of those of participants in the low-pressure condition.
In summary, there are quite a few areas of potential future research that this dissertation has revealed, and it will no doubt be rewarding to develop the temporal estimation paradigm, and note what results it generates, both intra-linguistically and cross-linguistically.

7.6 A Final Thought

For the past fifteen years, simulation semantics has been, for the most part, about lexical meaning. This dissertation, however, embraces (and, verily, confirms) the fact that there is much more to simulation-based representation than simple lexical meaning. And this is, of course, a good thing. The intricacies and complexities of a cognitive-functionalist model of language do not stifle simulation semantics or its reported findings. Rather, they enliven it with richness and varicolored intricacy that will insure a robust and wide range of research for many years to come.
REFERENCES


### Appendix 1: Critical Stimuli Used in Experiment One

<table>
<thead>
<tr>
<th>Sentence</th>
<th>Direction</th>
<th>Grammatical Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>Laura golpeaba la almohada.</td>
<td>Away</td>
<td>Imperfect</td>
</tr>
<tr>
<td>Ángel tiraba la naranja.</td>
<td>Away</td>
<td>Imperfect</td>
</tr>
<tr>
<td>Pilar rechazaba la cerveza.</td>
<td>Away</td>
<td>Imperfect</td>
</tr>
<tr>
<td>Pedro arrojaba la pelota.</td>
<td>Away</td>
<td>Imperfect</td>
</tr>
<tr>
<td>Rosa lanzaba el balón.</td>
<td>Away</td>
<td>Imperfect</td>
</tr>
<tr>
<td>Marta empujaba la puerta.</td>
<td>Away</td>
<td>Imperfect</td>
</tr>
<tr>
<td>Simón bebía la cerveza.</td>
<td>Toward</td>
<td>Imperfect</td>
</tr>
<tr>
<td>Victor recogía la almohada.</td>
<td>Toward</td>
<td>Imperfect</td>
</tr>
<tr>
<td>Inés comía la naranja.</td>
<td>Toward</td>
<td>Imperfect</td>
</tr>
<tr>
<td>Berto atrapaba el balón.</td>
<td>Toward</td>
<td>Imperfect</td>
</tr>
<tr>
<td>Pablo jalaba la puerta.</td>
<td>Toward</td>
<td>Imperfect</td>
</tr>
<tr>
<td>Ramón tomaba la pelota.</td>
<td>Toward</td>
<td>Imperfect</td>
</tr>
<tr>
<td>Marco va a arrojar la pelota.</td>
<td>Away</td>
<td>Morphological Future</td>
</tr>
<tr>
<td>Luis lanzará el balón.</td>
<td>Away</td>
<td>Morphological Future</td>
</tr>
<tr>
<td>José empujará la puerta.</td>
<td>Away</td>
<td>Morphological Future</td>
</tr>
<tr>
<td>Inés arrojará la pelota.</td>
<td>Away</td>
<td>Morphological Future</td>
</tr>
<tr>
<td>Hugo golpeará la almohada.</td>
<td>Away</td>
<td>Morphological Future</td>
</tr>
<tr>
<td>Norma rechazará la cerveza.</td>
<td>Away</td>
<td>Morphological Future</td>
</tr>
<tr>
<td>Eva tirará la naranja.</td>
<td>Away</td>
<td>Morphological Future</td>
</tr>
<tr>
<td>Sara jalará la puerta.</td>
<td>Toward</td>
<td>Morphological Future</td>
</tr>
<tr>
<td>Laura atrapará el balón.</td>
<td>Toward</td>
<td>Morphological Future</td>
</tr>
<tr>
<td>César tomará la pelota.</td>
<td>Toward</td>
<td>Morphological Future</td>
</tr>
<tr>
<td>Inés va a comer la naranja.</td>
<td>Toward</td>
<td>Morphological Future</td>
</tr>
<tr>
<td>Moises recogerá la almohada.</td>
<td>Toward</td>
<td>Morphological Future</td>
</tr>
<tr>
<td>Jorge beberá la cerveza.</td>
<td>Toward</td>
<td>Morphological Future</td>
</tr>
<tr>
<td>Carmen comerá la naranja.</td>
<td>Toward</td>
<td>Morphological Future</td>
</tr>
<tr>
<td>Berto va a golpear la almohada.</td>
<td>Away</td>
<td>Periphrastic Future</td>
</tr>
<tr>
<td>Ruben va a empujar la puerta.</td>
<td>Away</td>
<td>Periphrastic Future</td>
</tr>
<tr>
<td>Simón va a rechazar la cerveza.</td>
<td>Away</td>
<td>Periphrastic Future</td>
</tr>
<tr>
<td>Carla va a lanzar el balón.</td>
<td>Away</td>
<td>Periphrastic Future</td>
</tr>
<tr>
<td>Marta va a tirar la naranja.</td>
<td>Away</td>
<td>Periphrastic Future</td>
</tr>
<tr>
<td>Ruben va a tomar la pelota.</td>
<td>Toward</td>
<td>Periphrastic Future</td>
</tr>
<tr>
<td>Jorge va a jalar la puerta.</td>
<td>Toward</td>
<td>Periphrastic Future</td>
</tr>
<tr>
<td>Pilar va a recoger la almohada.</td>
<td>Toward</td>
<td>Periphrastic Future</td>
</tr>
<tr>
<td>Coco va a atrapar el balón.</td>
<td>Toward</td>
<td>Periphrastic Future</td>
</tr>
<tr>
<td>Andrés va a beber la cerveza.</td>
<td>Toward</td>
<td>Periphrastic Future</td>
</tr>
<tr>
<td>Sara rechaza la cerveza.</td>
<td>Away</td>
<td>Present</td>
</tr>
<tr>
<td>Hugo lanza el balón.</td>
<td>Away</td>
<td>Present</td>
</tr>
<tr>
<td>Moises tira la naranja.</td>
<td>Away</td>
<td>Present</td>
</tr>
<tr>
<td>Coco golpea la almohada.</td>
<td>Away</td>
<td>Present</td>
</tr>
</tbody>
</table>
Jorge empuja la puerta. Away Present
Rita arroja la pelota. Away Present
Lupe atrapa el balón. Toward Present
Lupe bebe la cerveza. Toward Present
Rita come la naranja. Toward Present
Carlos jala la puerta. Toward Present
Sara toma la pelota. Toward Present
Marco recoge la almohada. Toward Present
Norma está empujando la puerta. Away Present Progressive
Berto está lanzando el balón. Away Present Progressive
Elsa está tirando la naranja. Away Present Progressive
Carlos está arrojando la pelota. Away Present Progressive
Ramón está rechazando la cerveza. Away Present Progressive
Pablo está golpeando la almohada. Away Present Progressive
Jorge está comiendo la naranja. Toward Present Progressive
Ángel está jalando la puerta. Toward Present Progressive
Moises está atrapando el balón. Toward Present Progressive
Luis está bebiendo la cerveza. Toward Present Progressive
Víctor está tomando la pelota. Toward Present Progressive
Sara está recogiendo la almohada. Toward Present Progressive
Carla arrojó la pelota. Away Preterite
Berto golpeó la almohada. Away Preterite
Carlos rechazó la cerveza. Away Preterite
César tiró la naranja. Away Preterite
Andrés lanzó el balón. Away Preterite
Carmen empujó la puerta. Away Preterite
Pablo tomó la pelota. Toward Preterite
Rosa recogió la almohada. Toward Preterite
Marta bebió la cerveza. Toward Preterite
Carlos atrapó el balón. Toward Preterite
Eva comió la naranja. Toward Preterite
Eva jaló la puerta. Toward Preterite
Parece que Norma está rechazando la cerveza. Away Subordinate Clause / Indicative
Parece que Carla está arrojando la pelota. Away Subordinate Clause / Indicative
Parece que Jorge está lanzando el balón. Away Subordinate Clause / Indicative
Parece que Luis esté empujando la puerta. Away Subordinate Clause / Indicative
Parece que Marco está tirando la naranja. Away Subordinate Clause / Indicative
Parece que José está golpeando la almohada. Away Subordinate Clause / Indicative
Parece que Elsa está empujando la puerta. Away Subordinate Clause / Indicative
Parece que Sara está tomando la pelota. Toward Subordinate Clause / Indicative
Parece que Ramón está bebiendo la cerveza. Toward Subordinate Clause / Indicative
Parece que Víctor está atrapando el balón. Toward Subordinate Clause / Indicative
Parece que Luis está jalando la puerta. Toward Subordinate Clause / Indicative
<table>
<thead>
<tr>
<th>Sentence</th>
<th>Direction</th>
<th>Subordinate Clause Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parece que Simón está recogiendo la almohada.</td>
<td>Toward</td>
<td>Subordinate Clause / Indicative</td>
</tr>
<tr>
<td>Parece que Jorge está comiendo la naranja.</td>
<td>Toward</td>
<td>Subordinate Clause / Indicative</td>
</tr>
<tr>
<td>Parece que Jorge esté tirando la naranja.</td>
<td>Away</td>
<td>Subordinate Clause / Subjunctive</td>
</tr>
<tr>
<td>Parece que Simón esté golpeando la almohada.</td>
<td>Away</td>
<td>Subordinate Clause / Subjunctive</td>
</tr>
<tr>
<td>Parece que Ramón esté rechazando la cerveza.</td>
<td>Away</td>
<td>Subordinate Clause / Subjunctive</td>
</tr>
<tr>
<td>Parece que Sara esté arrojando la pelota.</td>
<td>Away</td>
<td>Subordinate Clause / Subjunctive</td>
</tr>
<tr>
<td>Parece que Víctor esté lanzando el balón.</td>
<td>Away</td>
<td>Subordinate Clause / Subjunctive</td>
</tr>
<tr>
<td>Parece que José esté recogiendo la almohada.</td>
<td>Toward</td>
<td>Subordinate Clause / Subjunctive</td>
</tr>
<tr>
<td>Parece que Marco esté comiendo la naranja.</td>
<td>Toward</td>
<td>Subordinate Clause / Subjunctive</td>
</tr>
<tr>
<td>Parece que Elsa esté jalando la puerta.</td>
<td>Toward</td>
<td>Subordinate Clause / Subjunctive</td>
</tr>
<tr>
<td>Parece que Carla esté tomando la pelota.</td>
<td>Toward</td>
<td>Subordinate Clause / Subjunctive</td>
</tr>
<tr>
<td>Parece que Berto esté atrapando el balón.</td>
<td>Toward</td>
<td>Subordinate Clause / Subjunctive</td>
</tr>
<tr>
<td>Parece que Norma esté bebiendo la cerveza.</td>
<td>Toward</td>
<td>Subordinate Clause / Subjunctive</td>
</tr>
</tbody>
</table>
## Appendix 2: Critical Stimuli Used in Experiment Two

<table>
<thead>
<tr>
<th>Sentence</th>
<th>Tone Length</th>
<th>Condition</th>
<th>Verb Frequency (if Applicable)</th>
</tr>
</thead>
<tbody>
<tr>
<td>The man drove to Michigan often.</td>
<td>1.25</td>
<td>Habitual Event</td>
<td></td>
</tr>
<tr>
<td>The girl attended class often.</td>
<td>1.25</td>
<td>Habitual Event</td>
<td></td>
</tr>
<tr>
<td>The cat hissed often.</td>
<td>1.25</td>
<td>Habitual Event</td>
<td></td>
</tr>
<tr>
<td>The student overslept often.</td>
<td>1.25</td>
<td>Habitual Event</td>
<td></td>
</tr>
<tr>
<td>The child got lost often.</td>
<td>1.25</td>
<td>Habitual Event</td>
<td></td>
</tr>
<tr>
<td>The dog escaped the yard often.</td>
<td>1.25</td>
<td>Habitual Event</td>
<td></td>
</tr>
<tr>
<td>The woman traveled often.</td>
<td>1.25</td>
<td>Habitual Event</td>
<td></td>
</tr>
<tr>
<td>The child got lost once.</td>
<td>1.25</td>
<td>Single Event</td>
<td></td>
</tr>
<tr>
<td>The woman traveled once.</td>
<td>1.25</td>
<td>Single Event</td>
<td></td>
</tr>
<tr>
<td>The dog escaped the yard once.</td>
<td>1.25</td>
<td>Single Event</td>
<td></td>
</tr>
<tr>
<td>The cat hissed once.</td>
<td>1.25</td>
<td>Single Event</td>
<td></td>
</tr>
<tr>
<td>The man drove to Michigan once.</td>
<td>1.25</td>
<td>Single Event</td>
<td></td>
</tr>
<tr>
<td>The student overslept once.</td>
<td>1.25</td>
<td>Single Event</td>
<td></td>
</tr>
<tr>
<td>The girl attended class once.</td>
<td>1.25</td>
<td>Single Event</td>
<td></td>
</tr>
<tr>
<td>The carpenter built a chair.</td>
<td>2.25</td>
<td>Non-Stativized Event</td>
<td></td>
</tr>
<tr>
<td>The father dressed his children.</td>
<td>2.25</td>
<td>Non-Stativized Event</td>
<td></td>
</tr>
<tr>
<td>The janitor swept the floor.</td>
<td>2.25</td>
<td>Non-Stativized Event</td>
<td></td>
</tr>
<tr>
<td>The cook prepared the food.</td>
<td>2.25</td>
<td>Non-Stativized Event</td>
<td></td>
</tr>
<tr>
<td>The athlete ran a marathon.</td>
<td>2.25</td>
<td>Non-Stativized Event</td>
<td></td>
</tr>
<tr>
<td>The boy painted a picture.</td>
<td>2.25</td>
<td>Non-Stativized Event</td>
<td></td>
</tr>
<tr>
<td>The girl wrote a poem.</td>
<td>2.25</td>
<td>Non-Stativized Event</td>
<td></td>
</tr>
<tr>
<td>The girl was writing a poem.</td>
<td>2.25</td>
<td>Stativized Event</td>
<td></td>
</tr>
<tr>
<td>The boy was painting a picture.</td>
<td>2.25</td>
<td>Stativized Event</td>
<td></td>
</tr>
<tr>
<td>The cook was preparing the food.</td>
<td>2.25</td>
<td>Stativized Event</td>
<td></td>
</tr>
<tr>
<td>The athlete was running a marathon.</td>
<td>2.25</td>
<td>Stativized Event</td>
<td></td>
</tr>
<tr>
<td>The father was dressing his children.</td>
<td>2.25</td>
<td>Stativized Event</td>
<td></td>
</tr>
<tr>
<td>The carpenter was building a chair.</td>
<td>2.25</td>
<td>Stativized Event</td>
<td></td>
</tr>
<tr>
<td>The janitor was sweeping the floor.</td>
<td>2.25</td>
<td>Stativized Event</td>
<td></td>
</tr>
<tr>
<td>The nanny watched the children for one year.</td>
<td>1.5</td>
<td>Long Temporal Interval</td>
<td></td>
</tr>
<tr>
<td>The girl wrote in her journal for one year.</td>
<td>1.5</td>
<td>Long Temporal Interval</td>
<td></td>
</tr>
<tr>
<td>The boy practiced violin for one year.</td>
<td>1.5</td>
<td>Long Temporal Interval</td>
<td></td>
</tr>
<tr>
<td>The child played soccer for one year.</td>
<td>1.5</td>
<td>Long Temporal Interval</td>
<td></td>
</tr>
<tr>
<td>The actress rehearsed her lines for one year.</td>
<td>1.5</td>
<td>Long Temporal Interval</td>
<td></td>
</tr>
<tr>
<td>The man swam laps for one year.</td>
<td>1.5</td>
<td>Long Temporal Interval</td>
<td></td>
</tr>
<tr>
<td>The woman worked on the project for one year.</td>
<td>1.5</td>
<td>Long Temporal Interval</td>
<td></td>
</tr>
<tr>
<td>The actress rehearsed her lines for one hour.</td>
<td>1.5</td>
<td>Short Temporal Interval</td>
<td></td>
</tr>
<tr>
<td>The woman worked on the project for one hour.</td>
<td>1.5</td>
<td>Short Temporal Interval</td>
<td></td>
</tr>
<tr>
<td>The man swam laps for one hour.</td>
<td>1.5</td>
<td>Short Temporal Interval</td>
<td></td>
</tr>
<tr>
<td>The boy practiced violin for one hour.</td>
<td>1.5</td>
<td>Short Temporal Interval</td>
<td></td>
</tr>
<tr>
<td>The child played soccer for one hour.</td>
<td>1.5</td>
<td>Short Temporal Interval</td>
<td></td>
</tr>
<tr>
<td>Event</td>
<td>Duration</td>
<td>Type</td>
<td>Score</td>
</tr>
<tr>
<td>----------------------------------------------------------------------</td>
<td>----------</td>
<td>--------------------</td>
<td>-------</td>
</tr>
<tr>
<td>The girl wrote in her journal for one hour.</td>
<td>1.5</td>
<td>Short Temporal Interval</td>
<td></td>
</tr>
<tr>
<td>The nanny watched the children for one hour.</td>
<td>1.5</td>
<td>Short Temporal Interval</td>
<td></td>
</tr>
<tr>
<td>The lizard darted under the rock.</td>
<td>1.75</td>
<td>Fast-Speed Verb</td>
<td>658</td>
</tr>
<tr>
<td>The team charged across the field.</td>
<td>1.75</td>
<td>Fast-Speed Verb</td>
<td>832</td>
</tr>
<tr>
<td>The child dashed across the room.</td>
<td>1.75</td>
<td>Fast-Speed Verb</td>
<td>1013</td>
</tr>
<tr>
<td>The car zipped around the bend.</td>
<td>1.75</td>
<td>Fast-Speed Verb</td>
<td>826</td>
</tr>
<tr>
<td>The man sprinted down the street.</td>
<td>1.75</td>
<td>Fast-Speed Verb</td>
<td>858</td>
</tr>
<tr>
<td>The horse galloped past the barn.</td>
<td>1.75</td>
<td>Fast-Speed Verb</td>
<td>267</td>
</tr>
<tr>
<td>The man strolled down the street.</td>
<td>1.75</td>
<td>Slow-Speed Verb</td>
<td>1145</td>
</tr>
<tr>
<td>The horse trotted past the barn.</td>
<td>1.75</td>
<td>Slow-Speed Verb</td>
<td>477</td>
</tr>
<tr>
<td>The bird glided through the air.</td>
<td>1.75</td>
<td>Slow-Speed Verb</td>
<td>964</td>
</tr>
<tr>
<td>The team ambled across the field.</td>
<td>1.75</td>
<td>Slow-Speed Verb</td>
<td>182</td>
</tr>
<tr>
<td>The lizard crawled under the rock.</td>
<td>1.75</td>
<td>Slow-Speed Verb</td>
<td>2661</td>
</tr>
<tr>
<td>The car inched around the bend.</td>
<td>1.75</td>
<td>Slow-Speed Verb</td>
<td>570</td>
</tr>
<tr>
<td>The child toddled across the room.</td>
<td>1.75</td>
<td>Slow-Speed Verb</td>
<td>57</td>
</tr>
<tr>
<td>The dog was small.</td>
<td>2.5</td>
<td>Permanent State</td>
<td></td>
</tr>
<tr>
<td>The student was German.</td>
<td>2.5</td>
<td>Permanent State</td>
<td></td>
</tr>
<tr>
<td>The scientist was intelligent.</td>
<td>2.5</td>
<td>Permanent State</td>
<td></td>
</tr>
<tr>
<td>The woman was tall.</td>
<td>2.5</td>
<td>Permanent State</td>
<td></td>
</tr>
<tr>
<td>The cab driver was Mongolian.</td>
<td>2.5</td>
<td>Permanent State</td>
<td></td>
</tr>
<tr>
<td>The chef was French.</td>
<td>2.5</td>
<td>Permanent State</td>
<td></td>
</tr>
<tr>
<td>The athlete was talented.</td>
<td>2.5</td>
<td>Permanent State</td>
<td></td>
</tr>
<tr>
<td>The chef was angry.</td>
<td>2.5</td>
<td>Temporary State</td>
<td></td>
</tr>
<tr>
<td>The cab driver was irritated.</td>
<td>2.5</td>
<td>Temporary State</td>
<td></td>
</tr>
<tr>
<td>The athlete was exhausted.</td>
<td>2.5</td>
<td>Temporary State</td>
<td></td>
</tr>
<tr>
<td>The student was happy.</td>
<td>2.5</td>
<td>Temporary State</td>
<td></td>
</tr>
<tr>
<td>The woman was nervous.</td>
<td>2.5</td>
<td>Temporary State</td>
<td></td>
</tr>
<tr>
<td>The dog was tired.</td>
<td>2.5</td>
<td>Temporary State</td>
<td></td>
</tr>
<tr>
<td>The scientist was excited.</td>
<td>2.5</td>
<td>Temporary State</td>
<td></td>
</tr>
</tbody>
</table>
**Appendix 3: Critical Stimuli Used in Experiment Three**

<table>
<thead>
<tr>
<th>Sentence</th>
<th>Tone Length</th>
<th>Condition</th>
<th>Verb Frequency (if Applicable)</th>
</tr>
</thead>
<tbody>
<tr>
<td>The athlete swam across the lake.</td>
<td>1.5</td>
<td>Accomplishment</td>
<td></td>
</tr>
<tr>
<td>The designer painted a picture.</td>
<td>2</td>
<td>Accomplishment</td>
<td></td>
</tr>
<tr>
<td>The housekeeper cleaned the house.</td>
<td>1.75</td>
<td>Accomplishment</td>
<td></td>
</tr>
<tr>
<td>The musician sang a song.</td>
<td>2.25</td>
<td>Accomplishment</td>
<td></td>
</tr>
<tr>
<td>The old man jogged a mile.</td>
<td>1</td>
<td>Accomplishment</td>
<td></td>
</tr>
<tr>
<td>The olympian skied down the mountain.</td>
<td>1.25</td>
<td>Accomplishment</td>
<td></td>
</tr>
<tr>
<td>The teenager ate a hamburger.</td>
<td>2.5</td>
<td>Accomplishment</td>
<td></td>
</tr>
<tr>
<td>The burglar vanished.</td>
<td>1.25</td>
<td>Achievement</td>
<td>8256</td>
</tr>
<tr>
<td>The guard released the prisoner.</td>
<td>2</td>
<td>Achievement</td>
<td>40421</td>
</tr>
<tr>
<td>The lifeguard spotted a shark</td>
<td>1.5</td>
<td>Achievement</td>
<td>12950</td>
</tr>
<tr>
<td>The nanny recognized the child.</td>
<td>2.5</td>
<td>Achievement</td>
<td>57083</td>
</tr>
<tr>
<td>The politician succeeded.</td>
<td>1.75</td>
<td>Achievement</td>
<td>19728</td>
</tr>
<tr>
<td>The train departed.</td>
<td>1</td>
<td>Achievement</td>
<td>5703</td>
</tr>
<tr>
<td>The tutor noticed the mistake.</td>
<td>2.25</td>
<td>Achievement</td>
<td>44584</td>
</tr>
<tr>
<td>The athlete swam.</td>
<td>1.5</td>
<td>Activity</td>
<td>13019</td>
</tr>
<tr>
<td>The designer painted.</td>
<td>2</td>
<td>Activity</td>
<td>26546</td>
</tr>
<tr>
<td>The housekeeper cleaned.</td>
<td>1.75</td>
<td>Activity</td>
<td>21715</td>
</tr>
<tr>
<td>The musician sang.</td>
<td>2.25</td>
<td>Activity</td>
<td>47468</td>
</tr>
<tr>
<td>The old man jogged.</td>
<td>1</td>
<td>Activity</td>
<td>2359</td>
</tr>
<tr>
<td>The olympian skied.</td>
<td>1.25</td>
<td>Activity</td>
<td>5182</td>
</tr>
<tr>
<td>The teenager ate.</td>
<td>2.5</td>
<td>Activity</td>
<td>84477</td>
</tr>
<tr>
<td>The chef cooked the pasta.</td>
<td>1.5</td>
<td>Expected</td>
<td></td>
</tr>
<tr>
<td>The chef cooked the spaghetti.</td>
<td>1.5</td>
<td>Expected</td>
<td></td>
</tr>
<tr>
<td>The cook whisked the egg.</td>
<td>2.25</td>
<td>Expected</td>
<td></td>
</tr>
<tr>
<td>The cook whisked the ovum.</td>
<td>2.25</td>
<td>Expected</td>
<td></td>
</tr>
<tr>
<td>The father washed the dishes.</td>
<td>1</td>
<td>Expected</td>
<td></td>
</tr>
<tr>
<td>The father washed the plates.</td>
<td>1</td>
<td>Expected</td>
<td></td>
</tr>
<tr>
<td>The lady folded the garments.</td>
<td>1.25</td>
<td>Expected</td>
<td></td>
</tr>
<tr>
<td>The lady folded the laundry.</td>
<td>1.25</td>
<td>Expected</td>
<td></td>
</tr>
<tr>
<td>The maid shattered the urn.</td>
<td>2</td>
<td>Expected</td>
<td></td>
</tr>
<tr>
<td>The maid shattered the vase.</td>
<td>2</td>
<td>Expected</td>
<td></td>
</tr>
<tr>
<td>The messenger rode the bicycle.</td>
<td>1.75</td>
<td>Expected</td>
<td></td>
</tr>
<tr>
<td>The messenger rode the two-wheeler.</td>
<td>1.75</td>
<td>Expected</td>
<td></td>
</tr>
<tr>
<td>The supervisor chewed the gum.</td>
<td>2.5</td>
<td>Expected</td>
<td></td>
</tr>
<tr>
<td>The supervisor chewed the marshmallow.</td>
<td>2.5</td>
<td>Expected</td>
<td></td>
</tr>
<tr>
<td>The archer hit the target.</td>
<td>2.5</td>
<td>Semelfactive</td>
<td>84891</td>
</tr>
<tr>
<td>The housecat jumped.</td>
<td>2.25</td>
<td>Semelfactive</td>
<td>32456</td>
</tr>
<tr>
<td>The jock kicked the ball.</td>
<td>2</td>
<td>Semelfactive</td>
<td>23452</td>
</tr>
<tr>
<td>The little girl skipped.</td>
<td>1.25</td>
<td>Semelfactive</td>
<td>7413</td>
</tr>
<tr>
<td>The neon sign flashed.</td>
<td>1.5</td>
<td>Semelfactive</td>
<td>9974</td>
</tr>
<tr>
<td>-------------------------</td>
<td>------</td>
<td>---------------</td>
<td>------</td>
</tr>
<tr>
<td>The officer knocked on the door.</td>
<td>1.75</td>
<td>Semelfactive</td>
<td>19395</td>
</tr>
<tr>
<td>The young boy blinked.</td>
<td>1</td>
<td>Semelfactive</td>
<td>7035</td>
</tr>
<tr>
<td>The delegate minded her own business.</td>
<td>1.25</td>
<td>State</td>
<td>19928</td>
</tr>
<tr>
<td>The employee deserved a raise.</td>
<td>1</td>
<td>State</td>
<td>17243</td>
</tr>
<tr>
<td>The farmer preferred having sunshine.</td>
<td>1.75</td>
<td>State</td>
<td>23792</td>
</tr>
<tr>
<td>The orphan lacked clean clothing.</td>
<td>1.75</td>
<td>State</td>
<td>20826</td>
</tr>
<tr>
<td>The shoebox contained old photographs.</td>
<td>2.5</td>
<td>State</td>
<td>47087</td>
</tr>
<tr>
<td>The sophomore belonged to a club.</td>
<td>2</td>
<td>State</td>
<td>24779</td>
</tr>
<tr>
<td>The trader owned several suits.</td>
<td>2.25</td>
<td>State</td>
<td>39146</td>
</tr>
<tr>
<td>The chef cooked the sock.</td>
<td>1.5</td>
<td>Unexpected</td>
<td></td>
</tr>
<tr>
<td>The chef cooked the stocking.</td>
<td>1.5</td>
<td>Unexpected</td>
<td></td>
</tr>
<tr>
<td>The cook whisked the brew.</td>
<td>2.25</td>
<td>Unexpected</td>
<td></td>
</tr>
<tr>
<td>The cook whisked the coffee.</td>
<td>2.25</td>
<td>Unexpected</td>
<td></td>
</tr>
<tr>
<td>The father washed the flesh.</td>
<td>1</td>
<td>Unexpected</td>
<td></td>
</tr>
<tr>
<td>The father washed the meat.</td>
<td>1</td>
<td>Unexpected</td>
<td></td>
</tr>
<tr>
<td>The lady folded the mat.</td>
<td>1.25</td>
<td>Unexpected</td>
<td></td>
</tr>
<tr>
<td>The lady folded the rug.</td>
<td>1.25</td>
<td>Unexpected</td>
<td></td>
</tr>
<tr>
<td>The maid shattered the icicle.</td>
<td>2</td>
<td>Unexpected</td>
<td></td>
</tr>
<tr>
<td>The maid shattered the stalagmite.</td>
<td>2</td>
<td>Unexpected</td>
<td></td>
</tr>
<tr>
<td>The messenger rode the dragon.</td>
<td>1.75</td>
<td>Unexpected</td>
<td></td>
</tr>
<tr>
<td>The messenger rode the wyrm.</td>
<td>1.75</td>
<td>Unexpected</td>
<td></td>
</tr>
<tr>
<td>The supervisor chewed the foil.</td>
<td>2.5</td>
<td>Unexpected</td>
<td></td>
</tr>
<tr>
<td>The supervisor chewed the sheet metal.</td>
<td>2.5</td>
<td>Unexpected</td>
<td></td>
</tr>
</tbody>
</table>
## Appendix 4: Critical Stimuli Used in Experiment Four

<table>
<thead>
<tr>
<th>Slide Text</th>
<th>Construction</th>
<th>Verb Bias (Ditransitive/Oblique Goal)</th>
<th>Verb Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>You are flinging Amy the frisbee.</td>
<td>Ditransitive 0.8125</td>
<td>Caused Motion</td>
<td></td>
</tr>
<tr>
<td>You are flinging Amy the plate.</td>
<td>Ditransitive 0.8125</td>
<td>Caused Motion</td>
<td></td>
</tr>
<tr>
<td>You are flipping Cathy the card.</td>
<td>Ditransitive 0.666666667</td>
<td>Caused Motion</td>
<td></td>
</tr>
<tr>
<td>You are flipping Cathy the matchbook</td>
<td>Ditransitive 0.666666667</td>
<td>Caused Motion</td>
<td></td>
</tr>
<tr>
<td>You are lobbing Lucas the bottle.</td>
<td>Ditransitive 0.4</td>
<td>Caused Motion</td>
<td></td>
</tr>
<tr>
<td>You are lobbing Lucas the egg.</td>
<td>Ditransitive 0.4</td>
<td>Caused Motion</td>
<td></td>
</tr>
<tr>
<td>You are sending Jason the package.</td>
<td>Ditransitive 0.47160789</td>
<td>Caused Motion</td>
<td></td>
</tr>
<tr>
<td>You are sending Jason the shoebox.</td>
<td>Ditransitive 0.47160789</td>
<td>Caused Motion</td>
<td></td>
</tr>
<tr>
<td>You are sliding Jenna the glass.</td>
<td>Ditransitive 0.851851852</td>
<td>Caused Motion</td>
<td></td>
</tr>
<tr>
<td>You are sliding Jenna the tray.</td>
<td>Ditransitive 0.851851852</td>
<td>Caused Motion</td>
<td></td>
</tr>
<tr>
<td>You are throwing Julie the football.</td>
<td>Ditransitive 0.980769231</td>
<td>Caused Motion</td>
<td></td>
</tr>
<tr>
<td>You are throwing Julie the orange.</td>
<td>Ditransitive 0.980769231</td>
<td>Caused Motion</td>
<td></td>
</tr>
<tr>
<td>You are tossing Erin the apple.</td>
<td>Ditransitive 0.545454545</td>
<td>Caused Motion</td>
<td></td>
</tr>
<tr>
<td>You are tossing Erin the hat.</td>
<td>Ditransitive 0.545454545</td>
<td>Caused Motion</td>
<td></td>
</tr>
<tr>
<td>You are flinging the frisbee to Curtis.</td>
<td>Oblique Goal 0.8125</td>
<td>Caused Motion</td>
<td></td>
</tr>
<tr>
<td>You are flinging the plate to Curtis.</td>
<td>Oblique Goal 0.8125</td>
<td>Caused Motion</td>
<td></td>
</tr>
<tr>
<td>You are flipping the card to Caleb.</td>
<td>Oblique Goal 0.666666667</td>
<td>Caused Motion</td>
<td></td>
</tr>
<tr>
<td>You are flipping the matchbook to Caleb.</td>
<td>Oblique Goal 0.666666667</td>
<td>Caused Motion</td>
<td></td>
</tr>
<tr>
<td>You are lobbing the bottle to Sophie.</td>
<td>Oblique Goal 0.4</td>
<td>Caused Motion</td>
<td></td>
</tr>
<tr>
<td>You are lobbing the egg to Sophie.</td>
<td>Oblique Goal 0.4</td>
<td>Caused Motion</td>
<td></td>
</tr>
<tr>
<td>You are sending the package to Wanda.</td>
<td>Oblique Goal 0.47160789</td>
<td>Caused Motion</td>
<td></td>
</tr>
<tr>
<td>You are sending the shoebox to Wanda.</td>
<td>Oblique Goal 0.47160789</td>
<td>Caused Motion</td>
<td></td>
</tr>
<tr>
<td>You are sliding the glass to Nicole.</td>
<td>Oblique Goal 0.851851852</td>
<td>Caused Motion</td>
<td></td>
</tr>
<tr>
<td>You are sliding the tray to Nicole.</td>
<td>Oblique Goal 0.851851852</td>
<td>Caused Motion</td>
<td></td>
</tr>
<tr>
<td>You are throwing the football to Ashley.</td>
<td>Oblique Goal 0.980769231</td>
<td>Caused Motion</td>
<td></td>
</tr>
<tr>
<td>You are throwing the orange to Ashley.</td>
<td>Oblique Goal 0.980769231</td>
<td>Caused Motion</td>
<td></td>
</tr>
<tr>
<td>You are tossing the apple to Kathleen.</td>
<td>Oblique Goal 0.545454545</td>
<td>Caused Motion</td>
<td></td>
</tr>
<tr>
<td>You are tossing the hat to Kathleen.</td>
<td>Oblique Goal 0.545454545</td>
<td>Caused Motion</td>
<td></td>
</tr>
<tr>
<td>You are giving Carla the gift.</td>
<td>Ditransitive 2.980433055</td>
<td>Caused Possession</td>
<td></td>
</tr>
<tr>
<td>You are giving Carla the money.</td>
<td>Ditransitive 2.980433055</td>
<td>Caused Possession</td>
<td></td>
</tr>
<tr>
<td>You are handing Steven the ball.</td>
<td>Ditransitive 0.615735462</td>
<td>Caused Possession</td>
<td></td>
</tr>
<tr>
<td>You are handing Steven the shirt.</td>
<td>Ditransitive 0.615735462</td>
<td>Caused Possession</td>
<td></td>
</tr>
<tr>
<td>You are lending Sally the key.</td>
<td>Ditransitive 0.498108449</td>
<td>Caused Possession</td>
<td></td>
</tr>
<tr>
<td>You are lending Sally the phone.</td>
<td>Ditransitive 0.498108449</td>
<td>Caused Possession</td>
<td></td>
</tr>
<tr>
<td>You are loaning Carmen the notebook.</td>
<td>Ditransitive 1.352941176</td>
<td>Caused Possession</td>
<td></td>
</tr>
<tr>
<td>You are loaning Carmen the pencil.</td>
<td>Ditransitive 1.352941176</td>
<td>Caused Possession</td>
<td></td>
</tr>
<tr>
<td>You are offering Jackie the basket.</td>
<td>Ditransitive 1.649122807</td>
<td>Caused Possession</td>
<td></td>
</tr>
<tr>
<td>You are offering Jackie the cake.</td>
<td>Ditransitive 1.649122807</td>
<td>Caused Possession</td>
<td></td>
</tr>
<tr>
<td>You are passing Lisa the folder.</td>
<td>Ditransitive 0.494163424</td>
<td>Caused Possession</td>
<td></td>
</tr>
<tr>
<td>Action</td>
<td>Type</td>
<td>Value</td>
<td>Possession Type</td>
</tr>
<tr>
<td>--------------------------------------</td>
<td>------------</td>
<td>--------------</td>
<td>-----------------------</td>
</tr>
<tr>
<td>You are passing Lisa the scissors.</td>
<td>Ditransitive</td>
<td>0.494163424</td>
<td>Caused Possession</td>
</tr>
<tr>
<td>You are showing Patrick the picture.</td>
<td>Ditransitive</td>
<td>4.134706815</td>
<td>Caused Possession</td>
</tr>
<tr>
<td>You are showing Patrick the wallet.</td>
<td>Ditransitive</td>
<td>4.134706815</td>
<td>Caused Possession</td>
</tr>
<tr>
<td>You are giving the gift to Francis.</td>
<td>Oblique Goal</td>
<td>2.980433055</td>
<td>Caused Possession</td>
</tr>
<tr>
<td>You are giving the money to Francis.</td>
<td>Oblique Goal</td>
<td>2.980433055</td>
<td>Caused Possession</td>
</tr>
<tr>
<td>You are handing the ball to Walter.</td>
<td>Oblique Goal</td>
<td>0.615735462</td>
<td>Caused Possession</td>
</tr>
<tr>
<td>You are handing the shirt to Walter.</td>
<td>Oblique Goal</td>
<td>0.615735462</td>
<td>Caused Possession</td>
</tr>
<tr>
<td>You are lending the key to Ernie.</td>
<td>Oblique Goal</td>
<td>0.498108449</td>
<td>Caused Possession</td>
</tr>
<tr>
<td>You are lending the phone to Ernie.</td>
<td>Oblique Goal</td>
<td>0.498108449</td>
<td>Caused Possession</td>
</tr>
<tr>
<td>You are loaning the notebook to Laurie.</td>
<td>Oblique Goal</td>
<td>1.352941176</td>
<td>Caused Possession</td>
</tr>
<tr>
<td>You are loaning the pencil to Laurie.</td>
<td>Oblique Goal</td>
<td>1.352941176</td>
<td>Caused Possession</td>
</tr>
<tr>
<td>You are offering the basket to Paula.</td>
<td>Oblique Goal</td>
<td>1.649122807</td>
<td>Caused Possession</td>
</tr>
<tr>
<td>You are offering the cake to Paula.</td>
<td>Oblique Goal</td>
<td>1.649122807</td>
<td>Caused Possession</td>
</tr>
<tr>
<td>You are passing the folder to Gerald.</td>
<td>Oblique Goal</td>
<td>0.494163424</td>
<td>Caused Possession</td>
</tr>
<tr>
<td>You are passing the scissors to Gerald.</td>
<td>Oblique Goal</td>
<td>0.494163424</td>
<td>Caused Possession</td>
</tr>
<tr>
<td>You are showing the picture to Jerry.</td>
<td>Oblique Goal</td>
<td>4.134706815</td>
<td>Caused Possession</td>
</tr>
<tr>
<td>You are showing the wallet to Jerry.</td>
<td>Oblique Goal</td>
<td>4.134706815</td>
<td>Caused Possession</td>
</tr>
</tbody>
</table>